NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

A HUMAN FACTORS ANALYSIS AND CLASSIFICATION SYSTEM (HFACS) EXAMINATION OF COMMERCIAL VESSEL ACCIDENTS

by

Forrest J. Griggs

September 2012

Thesis Advisor: John K. Schmidt
Second Reader: Samuel E. Buttrey

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE A Human Factors Analysis and Classification System (HFACS) Examination of Commercial Vessel Accidents			**5. FUNDING NUMBERS**
6. AUTHOR Forrest J. Griggs			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000			**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ___N/A____.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited			**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Human error has been identified in an estimated 80% of all commercial and military maritime accidents. Crew sizes on commercial merchant ships are characteristically smaller than military vessels. Commercial merchant ships rely on automated technology in order to reduce crew sizes. Since next generation naval ship designs are leveraging automated technology in order to reduce manning, an examination of commercial ship accidents is warranted. Two independent raters coded 518 findings from 48 maritime mishap reports using the Department of Defense Human Factors Analysis and Classification System (HFACS) taxonomy. Inter-rater reliability was calculated using Cohen's Kappa and a final result of 0.72 was determined for HFACS Level I. HFACS analysis identified relationships among the HFACS levels and collision, allision, and grounding accidents. Logistic regression analysis identified six patterns stemming from latent conditions and active failures. This was used to develop a modified hazard analysis to identify how latent conditions aligned in the accident event chain, and to propose intervention measures. The research concluded that a maritime version of HFACS should be adopted to improve the reliability of classifying causal factors. Additionally, by employing human factors post-accident research the Navy may be able to develop possible intervention strategies for the fleet.

14. SUBJECT TERMS Human Systems Integration, Safety, Mishaps, Human Factors, Human Factors Analysis and Classification System (HFACS)			**15. NUMBER OF PAGES** 119
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 8-98) Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

A HUMAN FACTORS ANALYSIS AND CLASSIFICATION SYSTEM (HFACS) EXAMINATION OF COMMERCIAL VESSEL ACCIDENTS

Forrest J. Griggs
Lieutenant, United States Navy
B.S., Management Information Systems, Iowa State University, 2006
B.S., Logistics and Supply Chain Management, Iowa State University, 2006

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN HUMAN SYSTEMS INTEGRATION

from the

NAVAL POSTGRADUATE SCHOOL
September 2012

Author: Forrest J. Griggs

Approved by: John K. Schmidt
Thesis Advisor

Samuel E. Buttrey
Second Reader

Robert F. Dell
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Human error has been identified in an estimated 80% of all commercial and military maritime accidents. Crew sizes on commercial merchant ships are characteristically smaller than military vessels. Commercial merchant ships rely on automated technology in order to reduce crew sizes. Since next generation naval ship designs are leveraging automated technology in order to reduce manning, an examination of commercial ship accidents is warranted. Two independent raters coded 518 findings from 48 maritime mishap reports using the Department of Defense Human Factors Analysis and Classification System (HFACS) taxonomy. Inter-rater reliability was calculated using Cohen's Kappa and a final result of 0.72 was determined for HFACS Level I. HFACS analysis identified relationships among the HFACS levels and collision, allision, and grounding accidents. Logistic regression analysis identified six patterns stemming from latent conditions and active failures. This was used to develop a modified hazard analysis to identify how latent conditions aligned in the accident event chain, and to propose intervention measures. The research concluded that a maritime version of HFACS should be adopted to improve the reliability of classifying causal factors. Additionally, by employing human factors post-accident research the Navy may be able to develop possible intervention strategies for the fleet.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ARPA	Automatic Radar Plotting Aid
ATG	Afloat Training Group
BRM	Bridge Resource Management
C3I	Command, Control, Communications, and Information
CCG	Canadian Coast Guard
CNO	Chief of Naval Operations
CO	Commanding Officer
CIC	Combat Information Center
CD	Compact Disc
DAU	Defense Acquisition University
DoD	Department of Defense
DND	Department of National Defense
DOT	Department of Transportation
ECDIS	Electronic Chart Display and Information System
FY	Fiscal Year
GAO	Government Accountability Office
GEMS	Generic Error Modeling System
GPS	Global Positioning System
HFACS	Human Factors Analysis and Classification System
HFE	Human Factors Engineering
HFWG	Human Factors Working Group
HSI	Human Systems Integration
IMO	International Maritime Organization
IIC	Investigator-in-Charge
JSSC	Joint Services Safety Chiefs
JOOD	Junior Officer of the Deck
KSA	Knowledge, Skills, and Abilities
LCS	Littoral Combat Ship

LST	Landing Ship Tank
MPT	Manpower, Personnel, and Training
MAIB	Marine Accident Investigation Branch
MCA	Maritime and Coastguard Agency
MIT	Massachusetts Institute of Technology
MILPERS	Military and Personnel
NAVSEA	Naval Sea Systems Command
NRC	National Research Council
NTSB	National Transportation Safety Board
NPS	Naval Postgraduate School
NSC	Naval Safety Center
OOD	Officer of the Deck
OMP	Optimal Manning Project
ORM	Operational Risk Management
RAC	Risk Assessment Code
RoRo	Roll-on, Roll-off
SWOS	Surface Warfare Officer School
SCM	Swiss Cheese Model
TSB	Transportation Safety Board
U.S.	United States
U.K.	United Kingdom
UNITAS	United International Antisubmarine Warfare
USCG	United States Coast Guard

EXECUTIVE SUMMARY

An estimated 80% of all commercial and military maritime accidents are attributed to human error. Commercial merchant ships typically function with smaller crews than surface warships, relying less on manpower and more on automation. Given the movement toward reducing personnel and leveraging technology in the next generation of naval vessels, an examination of commercial ship accidents and their causes is warranted. Published maritime mishap reports involving collisions, allisions, and groundings from Canada's Transportation Safety Board, the United Kingdom's Marine Accident Investigation Branch, and the United States' National Transportation Safety Board, between 2006 and 2011, were studied.

Two independent raters coded 518 findings from 48 maritime mishap reports using the Department of Defense Human Factors Analysis and Classification System (HFACS) taxonomy. Inter-rater reliability was calculated using Cohen's Kappa to determine agreement between raters. Initially, a low Cohen's Kappa of 0.45 was obtained for Level I and 0.39 for Level II, but after some discussion among raters, a consensus was reached for 405 of the 518 findings for Level I and 357 findings for Level II. Cohen's Kappa was recalculated to establish a final score of 0.72 for Level I and 0.64 for Level II; both levels were classified as having "good" agreement. The two raters were unable to reach consensus on 113 findings in Level I, most often attributing their differences to ambiguous nanocode definitions.

Using logistic regression, the results were analyzed for patterns to determine relationships between the different levels of HFACS and collisions, allisions, and groundings. The first round of logistic regression analysis consisted of determining if there was a significant difference in the HFACS Level I categories among the three types of accidents. If an HFACS category was found to be significant, further analysis was conducted on the related subcategories. The second round of logistic regression analysis performed was

to determine if there was a significant difference in the HFACS Level II subcategories among the three types of accidents.

In the first round of logistic regression analysis, it was determined that the HFACS Level I category of Supervision was significant ($p = 0.05$) in predicting collisions versus non-collisions. Subsequent analysis of Supervision subcategories revealed no significant factors. The second round of logistic regression analysis revealed that the model was unstable, and therefore disregarded.

During the first round of logistic regression conducted for allisions versus non-allisions against HFACS Level I categories revealed that the Organization category was significant ($p = 0.02$). Further analysis on Organization subcategories revealed that the Resource/Acquisition Management subcategory was significant ($p = 0.03$). The second round of logistic regression analysis conducted for allisions revealed that two Organization subcategories, Violations ($p = 0.04$) and Organizational Processes ($p = 0.08$), were significant or near significant.

The first round of logistic regression analysis conducted for groundings versus non-groundings against HFACS Level I categories revealed that the Supervision category was significant ($p = 0.02$). Analysis on Supervision subcategories revealed that the Inadequate Supervision subcategory was slightly above significance ($p = 0.06$); therefore, it was singled out for further analysis. Subsequent analysis on the Inadequate Supervision subcategory only revealed it was significant ($p = 0.4$). The second round of logistic regression analysis conducted for groundings revealed three subcategories were significant. The Judgment and Decision-Making Errors ($p = 0.01$) and Inadequate Supervision ($p = 0.01$) subcategories were both considered significant. The subcategory of Skill-Based Errors was near significant ($p = 0.06$).

The predominant latent conditions leading to active failures were identified as patterns and then assessed in terms of effects, risk, and corrective measures in a modified hazard analysis. Patterns derived from significant active failures were evaluated for predominant latent conditions prevalent in the accident reports. Collision accidents were identified with one significant pattern derived from the latent condition of unsafe supervision. Allision accidents were identified with three significant patterns stemming from the latent conditions of the HFACS subcategories of Resource Management, Inadequate Supervisory Conditions, and Organizational Processes. Grounding accidents were identified with two significant patterns derived from active failures of Skill-Based, and Judgment and Decision-Making Errors. For Skill-Based Errors, two latent conditions in the HFACS Level II subcategories of Inadequate Supervision and Organizational Climate appeared to be significant. Judgment and Decision-Making Errors were frequently associated with the latent conditions in HFACS Level II subcategories of Inadequate Supervision, Organizational Processes, and Organizational Climate. Mitigating actions and interventions currently used by the Navy were recommended for all patterns with associated latent conditions. When corrective measures were applied, likelihood values were decreased by one level, thereby reducing the overall Risk Assessment Value.

The research concluded that the HFACS taxonomy is useful for analyzing maritime mishaps, but that the lack of nanocodes for the surface community made coding difficult for raters. A maritime version of HFACS should be adopted to improve the reliability of classifying causal factors. It is also necessary that raters conducting analysis be provided concentrated HFACS training prior to coding findings to improve inter-rater reliability. Additionally, the Navy should consider further human factors research, using mishap and near-mishap data, to develop intervention strategies that could potentially reduce surface mishap rates in the future.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would never have been able to finish my thesis without the guidance of my committee members, and the loving support and patience of my family.

First and foremost, my utmost gratitude to my thesis advisor, CAPT John K. Schmidt, for his direction and steadfast devotion to his students.

I would also like to thank Dr. Samuel E. Buttrey for his patience, expertise, and encouragement.

Additionally, I would like to express my deepest gratitude to the HSI cohort. The support we gave each other through our postgraduate studies was essential to our success and, in return, we have gained lifelong friendships.

Finally, I would like to express my love and appreciation to my wife Linda, my sons, Gavin and Logan, and my daughter Makenna, who all waited patiently and continuously supported my endeavors.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. OVERVIEW

There has been an ongoing battle in the Navy's surface community to reduce the number of mishaps every year. Today's Navy is constrained by the threat of declining budgets, manning reductions, and sustained operational tempo. Consequently, this can make the way the Navy conducts business unaffordable (O'Rouke, 2011). In December 2011, Admiral Greenert, Chief of Naval Operations (CNO), wrote in a *Proceedings* article that "Our future fleet will remain ready, with the maintenance, weapons, personnel, and training it needs, although it may be smaller than today as a result of fiscal constraints" (Greenert, 2011, para. 2).

It is critical to preserve military resources and avoid warfighter capability gaps, which could aggravate the safety in the Fleet. Resources are needed to not only provide the military with the technology necessary to fight a war, but also with the training to use the technology. Modern weapons are complex to use; therefore, military personnel must train regularly to understand the capabilities, limitations, and operations of the platform or system (Atlantic Fleet Training and Testing, 2010). On May 30, 2007, Secretary of Defense (SECDEF) Robert Gates set a goal for the Department of Defense (DoD) to achieve a 75% accident reduction by 2008 (Gates, 2007). Mishaps directly cost the DoD approximately $3 billion per year, with indirect costs exceeding four times that amount. This accident reduction goal has not been attained has since been reinforced by the current SECDEF and reinforced by the Secretary of the Navy through the Department of the Navy's Safety Vision for 2012 (Maybus, 2009).

According to the Research of Innovative Technology Administration Bureau of Transportation Statistics, over 76% of maritime mishaps stem from human factors (Surface Warfare School, 2007). Recognizing the underlying role of human factors in mishaps is critical in distinctly categorizing mishaps. While most mishaps should be avoidable, each year proves that mishaps are inevitable. From 2000 to 2009, the Navy surface community had a Class-A mishap (accidents involving death or costs exceeding $2 million) 10-year average of 5.9 mishaps per year with no common trend identified (Commander, Naval Surface Forces, 2010).

In the past 10 years, the average hours that a ship is underway has increased by over 30% (Naval Sea Systems Command [NAVSEA], 2011). This operational tempo suggests a need to increase ship life cycles in the Fleet, as well as crew deployment lengths (O'Rouke, 2011). Since over 76% of reported maritime accidents are attributed to human error, an increase in operational tempo adds to the notion that more human errors can occur, which may lead to an increase in accidents. Recently, during a speech at the Surface Navy Association's annual conference, the Vice CNO, Admiral Mark Ferguson, conveyed that the Navy's goal of 313 ships may need to be reassessed, particularly since the earlier goal was based on the maritime defense needs in 2006 (Munoz, 2012). Furthermore, Admiral Greenert, CNO, testified before the Congress that the current Fleet size will be reduced in 2012 and will not return to today's current levels until 2017. He added that the 2013 budget submission has the Fleet reaching active ship levels of about 300 ships (Greenert, 2012).

The reduction in the defense budget confirms the need to reduce the amount of manpower in the Fleet. Of the Navy's Fiscal Year (FY) 2012 $161.4 billion budget, Military and Personnel (MILPERS) accounted for $46.6 billion, or 28.9% of the total budget (Assistant Secretary of the Navy, Financial Management and Comptroller, 2011). Consequently, MILPERS is predictably one of the first targeted areas to trim. The Navy's ability to operate a Fleet that is motivated and relevant means the workforce needs to have career-long tactical and strategic training (Greenert, 2011).

After providing an address on April 3, 2012 at the Naval Postgraduate School (NPS) in Monterey, California, Commander Kirk S. Lippold, the USS Cole (DDG 67) Commanding Officer (CO) when it suffered a terrorist attack in 2000, responded to an audience question regarding the Navy's optimal manning program. He stated: "How good is automation if you don't have power . . . if I had to deal with the 70% manning instead of the 94% manning that we had that day . . . I don't know if we could have done it!" The costs associated with a casualty due to a mishap, just like combat, means that the ship is out of the fight and resources such as manpower are being utilized to save the ship. CDR Lippold further stated that, "reducing the size of crews and the size of the fleet is the reality and may put ships and crews at jeopardy" (Lippold, 2012).

B. BACKGROUND

In the book, *Minding the Helm*, the National Research Council (NRC) observed that the traditional shipboard structure for command, control, communications, and information induces the potential for human error (National Research Council [NRC], 1994). Whether it is Navy ships or commercial vessels, human-systems interactions are constantly subjected to the human error chain. At first glance, there are some obvious differences between commercial vessels and U.S. Navy warships. Apparent visual differences are size, color, displacement, and draft, but a more in-depth look reveals similarities in the organizational culture and structure.

Navigation risk hazards are commonly mitigated by the use of automated technology such as Electronic Chart Display and Information System (ECDIS), Automatic Radar Plotting Aid (ARPA), or the commercial off-the-shelf system named FURUNO. The reliance on "dead-man" alarms (predetermined alerts on systems notifying operators of possible problems) may potentially reduce the overall big picture awareness for the Officer of the Deck (OOD), since the system is expected to alert them to any potential problems with navigation or shipping. Since it is necessary for OODs to roam around the bridge to receive various inputs from different systems and people, the availability of "dead man" alarms may provide a false sense of security. The OOD may become complacent in his or her duty of maintaining awareness of the "big picture," thereby reducing the effectiveness of his or her watchstanding performance (NRC, 1994).

All ships are comprised of numerous systems of systems, which must be integrated in order for the ship to properly function. Integrated systems are expected to reduce both accident risk and work load, thereby reducing crew numbers and costs (NRC, 1994). The crew is also integrated into the ship and must be managed as a highly valued resource (NRC, 1994). Commercial vessels (and warships) typically utilize a traditional organizational hierarchy for Command, Control, Communications, and Information (C3I) to conduct operational duties (NRC, 1994). From the military aspect, this hierarchical structure is necessary to communicate orders from officers to the crew, which typically means top-down communications. This practice, however, inhibits the potential for bottom-up communication, which may be necessary to propagate vital information.

Both Navy ship Commanding Officers (COs) and commercial vessel Masters (Captains) are subjected to time-speed-distance problems to ensure that their ships are in the right place, at the right time, while considering mission requirements. OODs or Officers of the Watch can easily be inundated with information communicated from numerous sources. It is their job to sift through the information, determine the best course of action, and make decisions in a dynamic bridge environment. The possibility for active failures or latent conditions to occur on a bridge (whether Navy or commercial) is great, considering the amount of information that flows between watch-stations.

Another critical factor in effective performance is trust, in terms of the system and operator performance (NRC, 1994). In order to have safe and reliable operations, there is a need for operators to have a deep intuitive knowledge based on training and experience, as well as trust while leveraging technologically advanced systems (NRC, 1994). The combination of technological, environmental, and organizational factors influence the way people perform (Rothblum et al., 2002). Automated technology is often designed without due regard to how the human operator needs to interact with the system, thereby impacting overall performance (Rothblum et al., 2002). Critical information may either not be displayed, or displayed in a way where the information is misinterpreted by the human (Rothblum et al., 2002).

The traditional organizational structure for C3I onboard ships can induce instead of reduce human error (NRC, 1994). With a traditional crew it may be apparent that more communication is needed in order to spread information or orders. Minimal crew size reduces the need for excessive communication, but, in turn, it may increase the training needed to operate systems and the workload for each watch-stander (NRC, 1994). The U.S. crew size on modern commercial vessels has decreased from a range of 30 to 40, to crews of approximately 20 to 30 per ship (NRC, 1994). Until recently, the Navy tried to use an Optimal Manning Project (OMP) to reduce the size of crews on certain platforms of ships in the Fleet to save money. OMP was intended not only as a cost-cutting measure, but also as a way to make the ships more efficient. In January 2011, CNO Admiral Greenert rescinded this notion due to the downward trend in mission readiness and the additional workload and training placed on crews (Schonberg, 2012). Today, instead of cutting a percentage of the crew, new designs for Navy ships are being built for an OMP-sized crew. This, however, necessitates the need for a higher reliance on automation and technology, along with resource management training for the crew.

Two U.S. Navy surface mishaps are cited extensively as case studies at the Surface Warfare Officer School in Newport, Rhode Island (Surface Warfare Officer School, 2007): The Landing Ship Tank (LST), *USS La Moure County* (LST 1194) and the Guided Missile Cruiser (CG), *USS Port Royal* (CG 73). Both ships had hard groundings and both mishaps involved human error, including failed resource management by leaders as well as an over-reliance on the use of readily available technology (Cole, 2009; McPherson, 2001). These two cases are clear examples of mishaps involving human error.

Many can recall images of the *USS Port Royal*, which ran aground on a coral reef just off the island of Oahu. The ship was conducting sea trials after an extensive yard period in Pearl Harbor (Kakesako, 2011). The grounding provided a clear example of how human factors issues can develop into a chain of errors resulting in an accident. The findings from the Navy Safety Investigation Board revealed that the CO was deficient on sleep; the ship's navigation system was mismanaged; some ship equipment was inoperable or faulty; and the bridge team was inexperienced (Cole, 2009). The cost to repair the coral reef and settle with the state of Hawaii was $8.5 million and the completed repairs to *USS Port Royal* were $40 million (Kakesako, 2011).

The *USS La Moure County* ran aground while conducting an amphibious assault exercise off the coast of Chile (McPherson, 2001). Improper use and reliance on Global Positioning System (GPS) technology, poor Bridge Resource Management (BRM), and failure to take into account the Combat Information Center's recommendations were just some of the links in the human error chain (McPherson, 2001). Since the cost to repair the ship was too great, it was struck from the Fleet, and on July 10 2001, it was towed into deep water and sunk as part of an annual United States-South American allied exercise named United International Antisubmarine Warfare (UNITAS) (NAVSOURCE.org, 2012).

Currently, the Bureau of Transportation Statistics estimates that 76% of civilian maritime mishap cases involve some form of human error (Surface Warfare Officer School, 2001). This is denoted in one classic incident involving the oil tanker *Exxon Valdez* and another, more recent, incident involving the cruise ship *Costa Concordia*. Both ship accidents were heavily publicized by the international news media and scrutinized for obvious human errors (Alaska Oil Spill Commission, 1990; British Broadcasting Corporation [BBC], 2012).

The grounding of the *Exxon Valdez* in 1989 provided insight into the devastating consequences of human error and the impact of having a reduced crew. The catastrophic accident turned into the worst environmental disaster at that time since Three Mile Island (Alaska Oil Spill Commission, 1990). Approximately 10.5 million gallons of crude oil spilled into Prince William Sound, which resulted in Exxon paying over $2 billion dollars to clean up the spill (Alaska Oil Spill Commission, 1990; Whitney, 1990). Excessive workload, crew fatigue, relaxed marine pilot standards, failure to properly maneuver the vessel, and failure to utilize technology in place on the bridge to recognize navigational hazards, all developed into a human error chain resulting in a commercial tanker accident (Whitney, 1990). The *Exxon Valdez* disaster was preventable not only by minimizing human error, but by adding more stringent safeguards, to including manning (Alaska Oil Spill Commission, 1990).

The BBC reported that the causes of the grounding and partial sinking of the *Costa Concordia* off the shores of Italy point to human error (BBC, 2012). Initial statements from the passengers and crew point to the Captain's inactions as being the cause. For some undetermined reason, the crew permitted the ship to go off its intended course and navigate through shoal waters (BBC, 2012). While the final investigation report has not been released yet, a statement by Costa Cruise Line boss Pier Luigi Foschi may foretell the outcome. He stated that "we need to acknowledge the facts and we cannot deny human error" (BBC, 2012, para. 10).

It is apparent by looking at these two civilian cases that human error patterns are prevalent just as with the mishaps seen involving U.S. Navy ships. The key point is that many of these ships, whether they be a cargo ship, tanker ship, or cruise ship, have minimal crew operating on the bridge and are exploiting the use of advanced technology. This is the current and highly visible trend in the new Navy ship designs being constructed today. Therefore, minimal manning ship designs may actually exacerbate the types of human errors seen in recent Navy ship mishaps, and it may also lead into other human error patterns due to the increasing reliance of technology on Navy ships.

C. OBJECTIVES

The purpose of this study was to use accident reports to conduct a systematic examination of commercial maritime navigation accidents in an attempt to understand existing human error patterns that may be pertinent to the Navy's surface Fleet. The intent was to identify human error patterns using the Human Factors Analysis and Classification System (HFACS) taxonomy adopted by the DoD. Once any human error patterns are identified, associated hazards are assessed in terms of risk to prioritize intervention development. The goal of this study was to provide recommendations for preventing human error in U.S. Navy ship operations, as the Navy moves toward leveraging technology to reduce ship crew size.

D. PROBLEM STATEMENT

Mishaps have been prevalent throughout the Fleet for a long period of time. The annual number of U.S. Navy afloat mishaps has decreased over the past decade, but so has the number of ships (Naval Safety Center, 2012). The number of ships in 2000 was 318, which has since been reduced to a current level of 285 in 2011 (Naval History & Heritage Command, 2011). The fact remains that mishaps will continue to occur and, with the reduced Fleet, each loss due to a mishap will become that much more critical.

An approach to isolating human error types that are relevant to the commercial shipping is to analyze maritime accidents and determine what occurred (Rothblum, 2002). Many of the analyses conducted on commercial vessels have shown that up to 80% of marine casualties had some form of human error (NRC, 1994). While these analyses have had some effect, particularly through awareness, it has not dramatically reduced the number of mishaps. Given that available resources are stretched for the U.S. Navy, it now has to redouble our efforts to reduce mishaps. This proposed study examined what types of human errors are occurring in current maritime accidents, to determine the root causal factors of maritime accidents and if patterns of human errors emerge.

The potential benefits of looking at human errors on commercial ships may be significant. This is especially true since there have been reductions to the defense budget. Reducing the amount of manpower on ships by increasing the use of automation and technology, while operational tempo increases and defense funding decreases, provides sufficient justification to apply the lessons found in the commercial vessel arena that could be applied to the U.S. Navy.

Lessons learned from commercial vessel accidents should be incorporated in the U.S. Navy's training and culture. Additionally, commercial vessel designs, technologies, and operating structure should be considered as part of the Analysis of Alternatives for Navy ship acquisitions. A hard look at the human factors aspect of why commercial vessels have accidents may be beneficial in reducing or preventing the Navy's surface mishaps in the future.

E. RESEARCH QUESTIONS

In order to identify human error patterns in commercial maritime accidents, this study systematically employs DoD HFACS to investigate publically available domestic and international accident data reports to address the following research questions.

- Can prevalent human error types and/or patterns be discerned in commercial maritime accidents?
- Is there a difference between prevalent human error types and/or patterns identified in groundings, collisions, and allisions?
- Given the presence of prevalent human error types and/or patterns, can relative risk be associated with those factors?
- What is the reliability of employing the HFACS taxonomy in classifying human error factors in commercial maritime accidents?

F. SCOPE AND LIMITATIONS

There is limited access to the Naval Safety Center's afloat mishap data concerning human error data analysis. This study is currently limited to commercial maritime accident reports that are available publicly via the web. This study is limited in our ability to gain access to Naval Safety Center afloat mishap data and, therefore, is restricted to reported domestic and international commercial marine accidents. This study will not include any mishaps that involve suicide as a leading cause of the mishap. Additionally, the extent of this research is primarily focused on merchant vessels.

G. HSI

> HSI is the technical process that integrates the disciplines of human factors engineering, manpower (number of people/workload on people), personnel (knowledge and skill requirements), training, habitability, personnel survivability, safety, and occupational health hazards concerns into the Systems Engineering of a material system to ensure safe, effective performance and maintainability. (Galdorisi & Truver, 2011, sect. 7)

Several of the HSI domains are involved in some manner in any mishap involving human error. Of these domains, this study will center its focus on Human Factors Engineering (HFE), and Manpower, Personnel, and Training (MPT).

1. HFE

HFE is primarily concerned with designing human-machine interfaces consistent with the physical, cognitive, and sensory abilities of the user population (Defense Acquisition University [DAU], 2009). HFE is one of the HSI domains that focus on the trade-offs among safety and the "ilities" such as system reliability, operability, usability, and maintainability. Regarding maritime accidents, HFE looks at many areas including lighting, visibility, noise, vibration, human fatigue, automated technology, command, control, communication, and human-machine environments (Calhoun, 2006). Many ships' automated bridge systems are often designed with the notion that they are "fail-safe," meaning the actions or inactions of bridge watchstanders will be identified and either corrected by the system itself, or at the very least alert the human operator of the impending error. The idea that any system is foolproof may be a misnomer and give the human operator a false sense of comfort. Additionally on Navy ships, bridge watch teams often rely on watchstanders in Combat Information Center (CIC) to provide protection for any of their errors. Humans, in many ways, are ingenious and will find ways to counter the "fail-safe" system whether done intentionally or unintentionally.

2. MPT

Manpower represents the number of personnel or positions required to perform a specific task (DAU, 2009). Plans to reduce the overall force structure of the Navy suggest there will be fewer people and fewer ships. Unfortunately, the mission requirements of the Navy have not diminished. This has a direct impact on the Navy's manpower, personnel, and training. New ship designs are dramatically reducing the number of crew onboard, while aging ships have recently reduced crew size to save funding. While the optimal manning experiment has since been retracted, some ships are now being decommissioned earlier than scheduled due to the pending $450 billion FY 2013 defense budget cut (Cavas, 2012). This new requirement increases the amount of tasking per ship, potentially overstressing crews.

Personnel considers the human aptitudes; knowledge, skills, and abilities (KSAs); and experience levels needed to perform job tasks (DAU, 2009). To minimize the effect of overtasking a specific sailor or crew, ships are being designed and built with more technological advances than ever before. This means that greater emphasis should be placed on human-system interaction in order for crews to understand and operate these systems (Dobie, 2003). The Navy recently reduced the number of qualified lookouts from two watchstanders to just one watchstander (U.S. Government Accountability Office [GAO], 2010). This was achieved by more effectively using the technology onboard ships and ensuring that personnel were properly qualified (GAO, 2010). The reduction in manpower for watch stations meant personnel needs were lessened onboard the Navy's cruisers and destroyers (GAO, 2010).

Training consists of the processes, procedures, and techniques used to train and qualify personnel, both military and civilian personnel to operate and maintain a system proficiently (DAU, 2009). There will be an increased training cost per sailor to optimize performance, and an extended training period to allow the sailors to complete the training pipeline. Since the training time is significantly longer for sailors assigned to a Littoral Combat Ship (LCS), detailers will open billets 18 to 24 months ahead of time rather than the six to nine months currently used (Commander, Naval Surface Force Public Affairs, 2006).

Trade-offs among manpower, personnel, and training influence the drivers of cost, performance, and schedule in new ship designs. A study conducted at the Naval Postgraduate School (NPS) showed that there are direct correlations between both manning and system performance, and inverse correlation between manning and mishap rates (Lazzaretti, 2008). These are significant relationships to consider, especially since ships are now being designed with smaller crew sizes, thus increasing the possibility of human error. The revelation of latent errors may take years, while active errors may affect the ship on a larger scale since each individual has more responsibilities.

3. Human Systems Integration Trade-Off Analysis

As defined in Secretary of the Navy Instruction 5000.2E, "HSI is the integrated analysis, design, and assessment over the life-cycle of a system and associated support infrastructure MPT, HFE, survivability, habitability, safety, and occupational health" (Department of the Navy, 2011). Trade-offs across the HSI domains are often difficult to quantify; therefore a qualitative description is currently used. Nevertheless, the goal of HSI analysis is to ensure that the requirements related to the HSI domains are satisfied within the constraints of the system's life-cycle cost, performance, and schedule (Holness, Shattuck, Winters, Pharmer, & White, 2011). In addition to understanding trade-offs, it is essential for HSI trade-off analysis to understand the boundaries of the problem in order to determine the "trade space." The trade space is the set of program and system parameters, attributes, and characteristics required to satisfy performance standards (Brantley, McFadden, & Davis, 2002). The term "trade space" has several connotations, but ultimately the trade space is dependent on the decision-maker's choices in regards to cost, schedule, and performance. Thus, understanding the problem, determining the trade space, and quantifying all possible trade-offs among the HSI domains will facilitate decision makers in the design and acquisition of a system, as well as operations involving the Navy fleet or the civilian sector.

Human factors issues such as sleep deprivation, the reliance on automated technology and extended working hours have persistently been noted in maritime accident reports. Fatigue can affect the overall performance of an individual, along with his or her ability to be attentive and remain alert (Calhoun, 2006). During periods of fatigue, there is an increased likelihood of diminished problem-solving ability, delayed reaction time, and increased risk taking. This typically results from sleep deprivation, poor sleep quality, stress, or physical/mental exertion (Calhoun, 2006). Fatigue in the captain and the crew was one of the factors that caused the *USS Port Royal* to run aground (Oceania Regional Response Team, 2009).

Another human factors issue is the use of automated technology onboard ships. Human errors are often the result of technologies that are incompatible with optimal human performance (Rothblum, 2002). Today's ship designs use a lot of technology in order to minimize manning levels (Calhoun, 2006). This can be considered a method to reduce the life-cycle costs of a ship, since increased technology means less people and less money. The reduction in crew sizes, however, does not come without some form of trade-off. Since the use of automation means a smaller crew, the crew is more susceptible to fatigue, stress, and task overload, which results in an increased risk for them and the ship (Calhoun, 2006). Additionally, trade-off considerations should be given to the likelihood of increased training requirements for operators and maintainers of a system. In general, human factors issues are difficult to resolve and engineers are usually compromised by economic pressures (Calhoun, 2006). The effects on other domains should be identified as early as possible and be quantifiable with regards to cost, performance, and schedule.

Regarding the domains of MPT, manning is considered to be a fixed number allotted for each ship of the Fleet. The current manning levels of ships has been decreased in order to reduce ship costs. Since 2001, the U.S. Navy has reduced manning levels to determine the optimal manning for a ship. Enlisted manning requirements throughout the Fleet were reduced by about 20% (GAO, 2010). The result of this manning decision is an indirect trade-off not using HSI practices. The standard workweek was increased from 67 to 70 hours, which can lead to crew fatigue and increase human errors (GAO, 2010; Naval History & Heritage Command, 2011). The true standard workweek, however, is actually about 80 hours once training and other factors are considered. The decrease in manning was offset by the increase in technology such as automated systems, and an increase in training and education for the sailors in order to utilize the automated systems. Other trade-off considerations are the financial costs associated with increasing automation to facilitate the loss of manpower; the increase in defense funding required for the training and education of sailors who will operate these automated systems; and the funding required in order to preserve the retention of sailors who may find the increased workload unappealing (Holness et al., 2011).

In order to reduce costs, minimizing the crew size onboard a warship to include the number of officers meant that a junior officer would be directly affected in terms of span and control and adequate manning to perform specific missions. To save training costs in 2003, the surface warfare community cut funding for the initial surface warfare school held in Newport, Rhode Island. All initial surface warfare training curricula were placed onto a compact disc (CD) "Surface Warfare Officer School (SWOS)-in-a-Box" and distributed to the ships in the Fleet for newly accessed officers to self-train (Shovlin, 2008). The impetus of the training was placed not only on the newly commissioned officer, but also on the more experienced department heads on the ships whose schedules were already fully saturated. Differences between the CD-based training and the traditional SWOS classroom environment were considered significant (Bowman, Crawford, & Mehay, 2008). The workload placed on the unqualified junior officer and department heads was considered too high and many senior Naval officers considered the training ineffective. The reduction of manpower on ships meant that officers needed to be better-trained surface warfare professionals. Instead, the trade-off was increased training costs due to the realization that newly implemented officers into the SWO community needed initial training. New schoolhouses had to be opened at the Afloat Training Groups (ATGs) and SWOS had been reinstated (Naval Surface Forces Public Affairs, 2011). It is plausible that the reduction in manpower, the deficiency of properly trained surface warfare junior officers, and the workload placed on junior officers and department heads placed ships in jeopardy, both in the present and in the future, as these insufficiently trained junior officers become department heads.

BRM is also challenged as the reduction of manning and personnel is placed on ships. During higher-risk evolutions such as entering or leaving port, crewmembers are asked to perform events beyond their normal scope of training. Aboard U.S. Navy frigates, junior Yeomen become bearing takers and bearing recorders for the Navigation team on the bridge, a task traditionally saved for trained Quartermasters who are essentially navigation specialists for the ship. A junior Yeoman may not understand the importance of reading off the three digits of a bearing to a physical navigation object (on land or in the water) through a sound-powered phone to the bearing recorder who, in turn, provides that information to the Navigation plotter. The Navigation plotter quickly takes that historic information and triangulates a position, or “fix,” onto a chart to determine where the ship *was* located as the bearing recorder initiated the sequence of the fix. This historical information is crucial to the upper levels of the operational chain of command on that ship, but the information chain may start with an inadequately trained junior Yeoman.

Bridge watch teams are often left without properly trained Junior OOD (JOOD) or Conning Officers as the ship conducts a harbor transit. The JOOD is, for all intents and purposes, the assistant to the OOD. The OOD is required to maintain the entire situation or “big picture” of the ship and apply it to the mission at hand. The immediate actions of the JOOD should reduce the workload of the

OOD, allowing him or her to maintain focus on the big picture; but, in the absence of a JOOD, at least an experienced one, the workload increases significantly. Situations such as this are becoming the new norm on ships and places a new requirement on BRM for OODs as they are required to do more with either fewer or less-trained personnel.

H. ORGANIZATION

This thesis is organized into five chapters. Chapter I presents an overview of human error involvement on recent accidents involving commercial vessels and attempts to apply this to both contemporary and classic Navy ship mishaps. Chapter II provides an understanding of human error and a description of the HFACS rating process. Chapter III describes the methodology of the HFACS rating process used on civilian maritime accident reports and the process used in conducting data analysis. Chapter IV presents the results of the HFACS rating process, HFACS categories that were found significant, and patterns identified, and a subsequent hazard analysis developed from the patterns. Chapter V offers conclusions and recommendations based on the findings.

II. LITERATURE REVIEW

A. OVERVIEW

This chapter provides a general overview of recent human factors studies involving surface vessels, human error, accident causation, and HFACS. The literature reviewed covered human factors issues, human error as causal factors (particularly involving safety and ship accidents), and renowned experts in the human error field. The literature consisted of thesis research reports, books written by renowned human error authors, and research published by government agencies. An in-depth review of the HFACS taxonomy was also conducted in order to gain a comprehensive understanding of the relation between Reason's Swiss Cheese Model (SCM) and HFACS. Furthermore, DoD HFACS materials were used for the purpose of this research.

B. MARITIME ACCIDENTS

A Massachusetts Institute of Technology (MIT) study examined causal factors contributing to ship groundings while transiting in and out of ports (Lin, 1998). It found that errors in tide forecasts did not have a significant effect as a risk factor for ship groundings while transiting in and out of port, but noted that nighttime transits were more hazardous than daytime transits (Lin, 1998). The study argued for future risk analysis, vis-à-vis human factors, as the most significant contribution to ship groundings in order to build a risk model for groundings (Lin, 1998). Complementary to this writing, the research at MIT found many papers indicating human error causal factors as the most significant reason for groundings, but very little human factors research has been conducted to support this notion (Lin, 1998). This is due largely in part to the lack of available historical data. This is not necessarily due to the inaccessibility of reports, but to the absence of an effort to record information from ship groundings (Lin, 1998).

A study conducted at NPS evaluated an earlier version of the HFACS taxonomy for applicability, reliability, and usefulness in a *post hoc* analysis of Naval Safety Center mishap data. In the HFACS study, it was found that the application of HFACS was relevant to Navy afloat Class “A” mishaps (i.e., mishaps that have fatalities or a total cost equal to or greater than $1 million) (Lacy, 1998). The HFACS taxonomy provided sufficient categories to classify mishap findings conducted in a *post hoc* data analysis of Naval Safety Center (NSC) surface accident reports. Over 92% (459 of 496) of all causal factors were classified in HFACS. While only one category was used per causal factor, it identified some causal factors that could be classified in more than one category (Lacy, 1998). The research concluded that HFACS was useful for identifying causal factor classification and supported strategies to prevent future mishaps (Lacy, 1998). The HFACS taxonomy has since been modified to classify all Department of Defense (DoD) operational and training mishaps (DoD, 2005).

In the NPS study, two Naval officers were used as raters to classify each report’s findings into the HFACS taxonomy and Cohen’s Kappa was used to assess inter-rater reliability (Lacy, 1998). After disagreements were resolved, from the data it was determined the HFACS taxonomy is useful for classifying human error causal factors for postaccident data analysis (Lacy, 1998). It was further recognized that U.S. Navy mishap reports, as currently written, are not conducive for human factors analysis and it was suggested that HFACS causal factor descriptions should be included within the database to better support analysis for identifying frequency of human error causal factors and patterns (Lacy, 1998).

An Human Factors Working Group (HFWG) report revealed that human error contributes to between 84% and 88% of tanker accidents; 79% of towing vessel groundings; 89% to 96% of collisions; 75% of allusions; and 75% of fires and explosions (Rothblum, 2002). These percentages show the importance of identifying precisely what the human error is in order to determine the appropriate intervention strategy for preventing future mishaps. Human factor issues such as fatigue, communication, BRM, automation, situational awareness, teamwork, decision-making and health and stress, are all factors that influence safety onboard maritime vessels.

A *Journal of Safety Research* article stated that, “Monitoring and modifying the human factors issues could contribute to maritime safety performance” (Hetherington, Flin, & Mearns, 2006, p. 401). The review recognizes the same issue that Lin’s research at MIT identified. That is, studies have been published regarding maritime safety, but no literature review has been conducted to “aggregate the causal factors within accidents in shipping and surmise current knowledge” (Hetherington et al., 2006, p. 401).

Clearly, numerous studies have revealed human error as a major component of ship accidents. There is an obvious safety element involved in the identification and classification of human error root causal factors in accidents. The International Maritime Organization (IMO) states, “Shipping is perhaps the most international of all the world’s great industries and one of the most dangerous” (Hetherington et al., 2006, p. 401). This statement emphasizes the importance of determining what causal factors were involved in maritime mishaps in order to identify patterns in human factors analysis and eventually develop interventions to prevent further human error resulting in accidents.

C. HUMAN ERROR

Human error has been identified in numerous writings as the root cause in both commercial and military maritime mishaps. The HFACS categorizations of human error have typically been completed through hindsight investigations, where the outcome of the mishap is known and the human is often attributed the blame (Salmon, Regan, & Johnston, 2005). Rasmussen considered the system itself in determining human error. If the system performs less satisfactorily because of a human act, then it is very likely human error (Rasmussen, 1986). He further defined errors as simply a difference between an actual state and a desired state (Rasmussen, 2003). Woods (2006) describes the labeling of “human error” as prejudicial. Utilizing the term “human error” hides much more than it reveals about how a system functions or malfunctions (Woods, Dekker, Cook, Johannesen, & Sarter, 2010). Another characterization of human error is the inappropriate human behavior that lowers levels of system effectiveness or safety, which may or may not result in an accident or injury (Wickens, Lee, Liu, & Gordon Becker, 2004).

This paper utilized Reason's definition of an error, which defines error as a symptom that reveals the presence of latent conditions in the system at large (Reason, 1997). Reason acknowledges that there is an overabundance of human error definitions without a universal definition of human error (Reason, 1990). Furthermore, there is no universal taxonomy to classify human error, for typically human error taxonomies are designed for a specific purpose, where no scheme is likely to satisfy all needs (Reason, 1990).

The term "error" is often used in a vague sense to easily describe actions or inactions that may have occurred in everyday events. In order to enable a more detailed examination, Reason further split errors into two main categories: errors and violations. Violations differ from errors because they are considered intentional acts, violations break known rules or procedures, whereas errors do not (Reason, 1990).

Violations are the intentional deliberate action(s) of a human operator (Reason, 1997). The person who committed the act made a conscious decision to deviate from the plan by his or her actions, thus breaking a specific rule or procedure, or defying the norms conveyed throughout the organization (Reason, 1997). Violations are often categorized into one of three areas: routine, optimizing, or exceptional. *Routine* violations are tolerated by the organization and tend to become habitual by nature (Shappell & Wiegmann, 2000). *Optimizing* violations occur when an individual seeks to elevate the importance of some goal other than safety (Maurino et al., 1995). *Exceptional* violations are actions that are considered an extreme departure from what is considered acceptable by either the organization or the operator. Reason describes this as a one-off breach of regulations dictated by unusual circumstances (Maurino, et al., 1995). While there are various categories of violations, and previous versions of HFACS had two subcategories of violations (routine and exceptional), this paper will align with the current DoD HFACS categories and not break violations down into any further subcategories.

In order to gain a better understanding of errors, Reason further categorized errors into three basic types (see Figure 1). He describes errors as skill-based slips and lapses, or rule-based and knowledge-based mistakes (Reason, 1990). Slips and lapses are better understood as failures in execution, while mistakes are considered failures in planning. Slips and lapses occur when there is a good plan in place and are considered unintentional. A slip is considered an incorrect action taken by a human operator (e.g., the wrong button was pressed). Lapses are failures in memory (e.g., an operator forgot to press the button). Mistakes occur when there is a failure to have a good plan in place, meaning a perfect execution of the plan would not have yielded the intended result. Plans which are poorly written, inappropriate for a particular mission, or too risky are very likely to have an operator make a mistake.

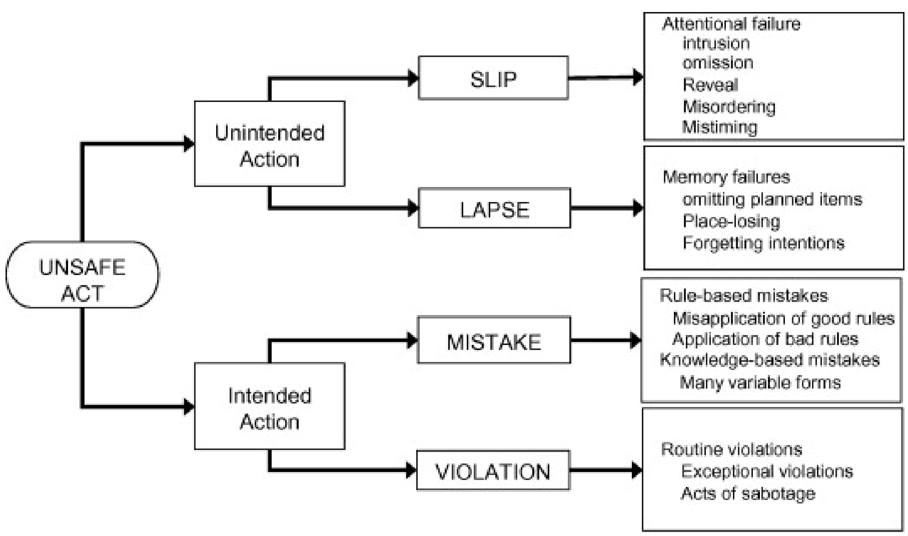

Figure 1. Slips, Lapses, and Mistakes (From Reason, 1994)

D. CAUSATION THEORIES

1. Knowledge-, Rule-, and Skill-Based Behavior

Reason's three basic error types of skill-based slips and lapses, and rule-based and knowledge-based mistakes can be correlated to Rasmussen's three performance levels (Reason, 1990). This is done by identifying the relationship between the three types of errors and the amount of cognitive processing or performance levels by the human in Rasmussen's levels of performance. The three levels of performance are knowledge-based, rule-based, and skill-based behaviors.

Skill-based behaviors represent sensory-motor performance during acts or activities that, following a statement of an intention, take place without conscious control as smooth, automated, and highly integrated patterns of behavior (Rasmussen, 1983). These can be common physical tasks, which require little cognitive thought. Flipping a switch upon hearing an alarm or turning a dial when an indicator reaches a certain parameter are examples. These actions have been conducted or practiced repeatedly and require virtually no monitoring—meaning the task becomes almost automatic.

Rule-based behaviors are like "if-then" statements (Airbus, 2005). If the problem is A, then do procedure B. This type of behavior is taught through formal training or by learning the job through the guidance of experienced workers, such as in an apprenticeship. The amount of conscious decision-making or cognitive thinking is higher than in rule-based behaviors, but less than in knowledge-based behaviors. Therefore, it is considered an intermediary performance level.

Reason's (1990) Generic Error Modeling System (GEMS) relates the three levels of performance in a taxonomy that relied on Rasmussen's three types of errors (see Figure 2). Knowledge-based behaviors are goal-based for this control of performance is at a much higher conceptual level than skill or rule-based behaviors. The goal is explicitly formulated and a mental model, which incorporates the functional properties of the environment, and predicts of the effects of the plan considered, is formed intuitively (Rasmussen, 1983). This level of performance is conducted in a virtual manner because there are no rules or training that have been derived to handle these types of situations.

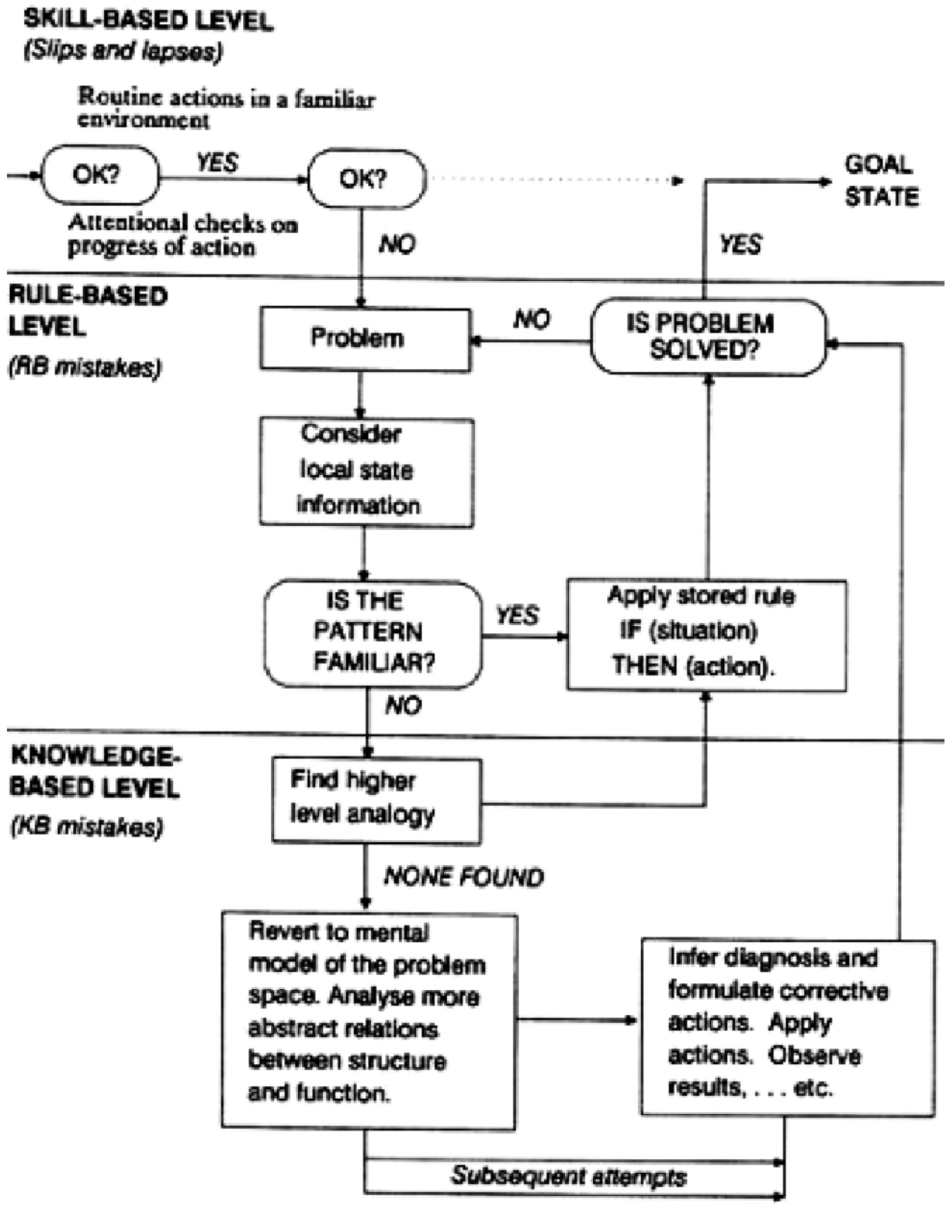

Figure 2. Generic Error Modeling System (From Reason, 1990)

2. Domino Theory

The core of the *Domino Theory*, developed by Herbert W. Heinrich who studied industrial safety in the early 1900s, is that accidents are a result of a sequence of events. This is analogous to dominos falling over in a line; as one falls over onto the other, it will cause a continuous chain of events resulting in an accident. Heinrich likened the dominos to unsafe conditions or unsafe acts, where their subsequent removal prevents a chain reaction from occurring, thus preventing an accident (Heinrich, 1941).

There are five dominos, or stages, defined by Heinrich, representing a figurative causal factor in accidents (see Figure 3). The first domino is *Social Environment and Ancestry*, which looks at how society or inheritance explains the personality of a worker. These people exhibit personality traits such as stubbornness, greed, or recklessness, which may have been developed from the social environment of the worker or essentially inherited (Heinrich, 1941). The second stage in the domino theory is the secondary personality traits that are gained by the worker due to their own personality. These traits, such as a bad temper or ignorance, and gained from society or inheritance, are essentially embedded with other character flaws and become contributors to unsafe acts or unsafe conditions. The third domino is the unsafe act or unsafe condition. This is the central reason why accidents occur in the workplace; thus, the subsequent removal of the unsafe act or unsafe condition will prevent an accident. The accident itself is the fourth domino. According to Heinrich (1941) an accident is just one factor in a sequence that results in an injury. The final domino is the injury to a person, which is a direct result of the accident. The combination of the first three dominos causes accidents, while the removal of any one of the dominos is the easiest way to prevent an accident from occurring.

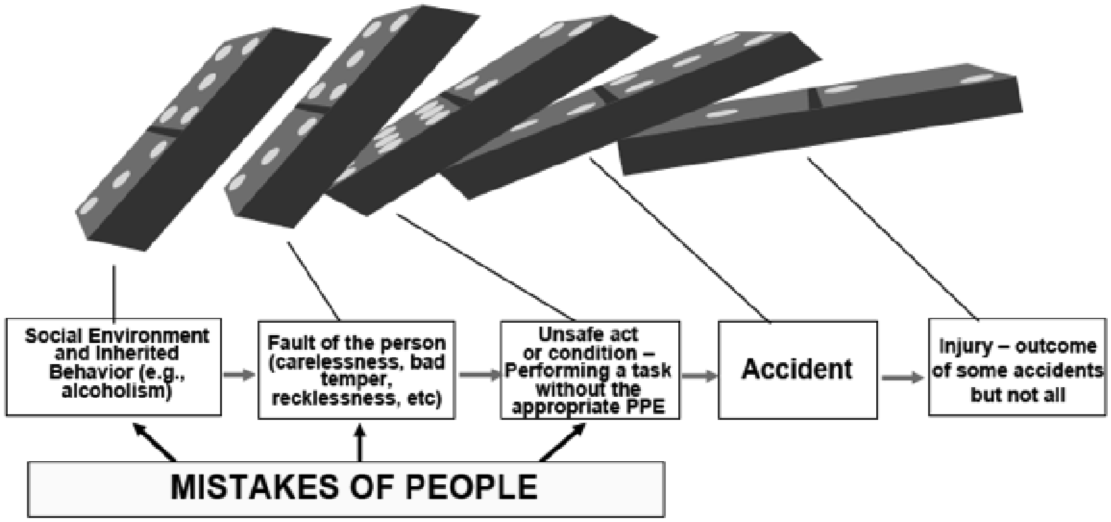

Figure 3. Heinrich Domino Theory (After Heinrich, 1941)

3. Swiss Cheese Model (SCM)

In 1990, Reason published his book, *Human Error*. The accident causation model he developed became better known (though not intentionally) as the Swiss Cheese Model. SCM simplifies a theory through a modest illustration of how a mishap or accident can occur in a system or organization. It places emphasis on the structure or hierarchy of an organization, and with the collaboration of human error. The model focuses on both organizational hierarchy and human error. Reason suggests that a mishap can occur due to the culmination of more than one human error occurring at each level of the SCM (Reason, 1990). The levels are essentially layers of defenses put in place to stop the chain of events, caused by human errors, leading to a mishap. His theory suggests that the ideal layered defense would not have any holes, thereby preventing any possible hazard leading to an accident. The reality of Reason's SCM shows that the layered defenses in depth are actually muddled with holes, which suggests that a mishap is virtually unavoidable if the holes were to line up.

This section offers an example of a sequence of events regarding failed defensive layers of U.S. Navy ship operation. It is analogous to the descriptions offered by Shappell and Wiegmann (2000), which were originally intended to gain insight about aviation defensive layers. An example of defensive layers failing is a possible scenario where the outcome of a mishap is that a U.S. Navy warship collides with a tanker vessel during a refueling operation. Prior to the mishap, fallible decisions by “Big Navy” or upper management may have resulted in aging ships being operated with less crew, but maintain the current operational tempo. The Captain may have been pressured to maintain mission readiness and to be in the right place, at the right time under unreasonable circumstances (Reason, 1990). The Officer of the Deck (OOD) may have been pressured to cope with a bridge team of less-trained personnel during a stressful refueling operation. Additionally, a less-than-ideal culture may have existed where little or no rest for the crew and officers was offered. This type of precondition is common for the Surface Warfare community, where minimal sleep is sometimes considered a rite of passage. The Conning Officer’s last conning order was misinterpreted by the Helmsman, which resulted in an unsafe act, where the Helmsman steered the warship towards the oiler instead of away. The Helmsman, meanwhile, may have been distracted and was dealing with an unknown alarm on the ship’s control console just prior to the collision. It is all too common in Fleet operations and is shown purely to demonstrate the effect of the holes in defense lining up just right to allow a Class “A” maritime mishap to occur. This introduced, or at least generally defined, a universal concept that mishaps are the culmination of errors that have passed through holes (or gaps in defenses) in the “cheese,” which are the layers of defenses designed to stop mishaps from occurring.

4. Active Failures and Latent Conditions

The holes in the SCM are derived from one of two notions that the holes are either due to active failures or some other latent conditions (see Figure 4). Active failures, also known as unsafe acts, are errors that are typically made by the main operator(s) of a system. Ships’ crews or aircraft pilots are examples of main operator(s) of a system. This is considered to be at the *sharp* end of the system and can have the most direct impact in allowing the mishap to occur. Since the actions of these operators can have an immediate impact on the mishap, the acts committed by them are characterized as active failures (Reason, 1990). Latent failures are considered the complement of active failures. These failures are customarily at the *blunt* end of the system and mishaps from such failures may take years to come to fruition. These mishaps were most likely not caused by the slips, lapses, or mistakes of the front-line operator. Instead, the error may have been deep within the system, hidden inside the organization. Years of neglected maintenance of a system is an example directly related to maritime vessels representing hidden, unsafe actions. These unsafe actions are considered to be consequential to the mishap, instead of being labeled as the direct causal factor (Reason, 1990).

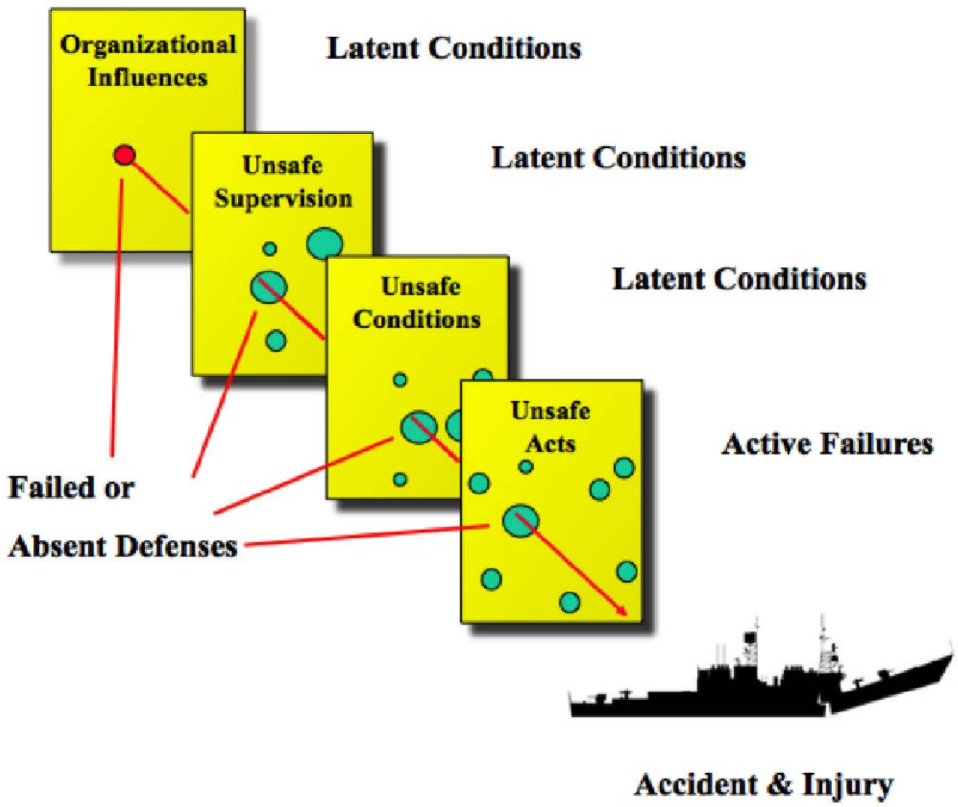

Figure 4. Reason's Swiss Cheese Model (From Naval Safety Center, 2012)

E. DOD HUMAN FACTORS ANALYSIS CLASSIFICATION SYSTEM

In an effort to operationalize Reason's Accident Causation theories and to be able to apply it, HFACS was developed. HFACS was originally developed for

Naval aviation, but it has since been directed for the application of surface fleet mishaps (Shappell & Wiegmann, 2000). HFACS has also been applied in the commercial sector as well as medical fields (HFACS, Inc., 2010).

HFACS draws directly upon Reason's SCM. Shappell and Wiegman (2000) describe Reason's SCM as an appealing approach to the genesis of human error. The fallibility of his model, however, is that it is just a general conception of how mishaps can occur and does not give detailed ways to apply it in real mishaps. The derivation of HFACS was a substantial attempt to classify human errors into a framework for studying aviation accidents. Reason's concept of active failures and latent conditions is the backbone of HFACS. Figure 5 shows the four levels of failure: (1) Unsafe Acts, (2) Preconditions for Unsafe Acts, (3) Unsafe Supervision, and (4) Organizational Influences. These are described in order to provide taxonomy to accident investigation (Shappell & Wiegmann, 2000). The original HFACS taxonomy was developed based on 300 Class A Naval Aviation flight mishaps, but it has been further cultivated into the DoD HFACS currently in use to investigate all operational mishaps.

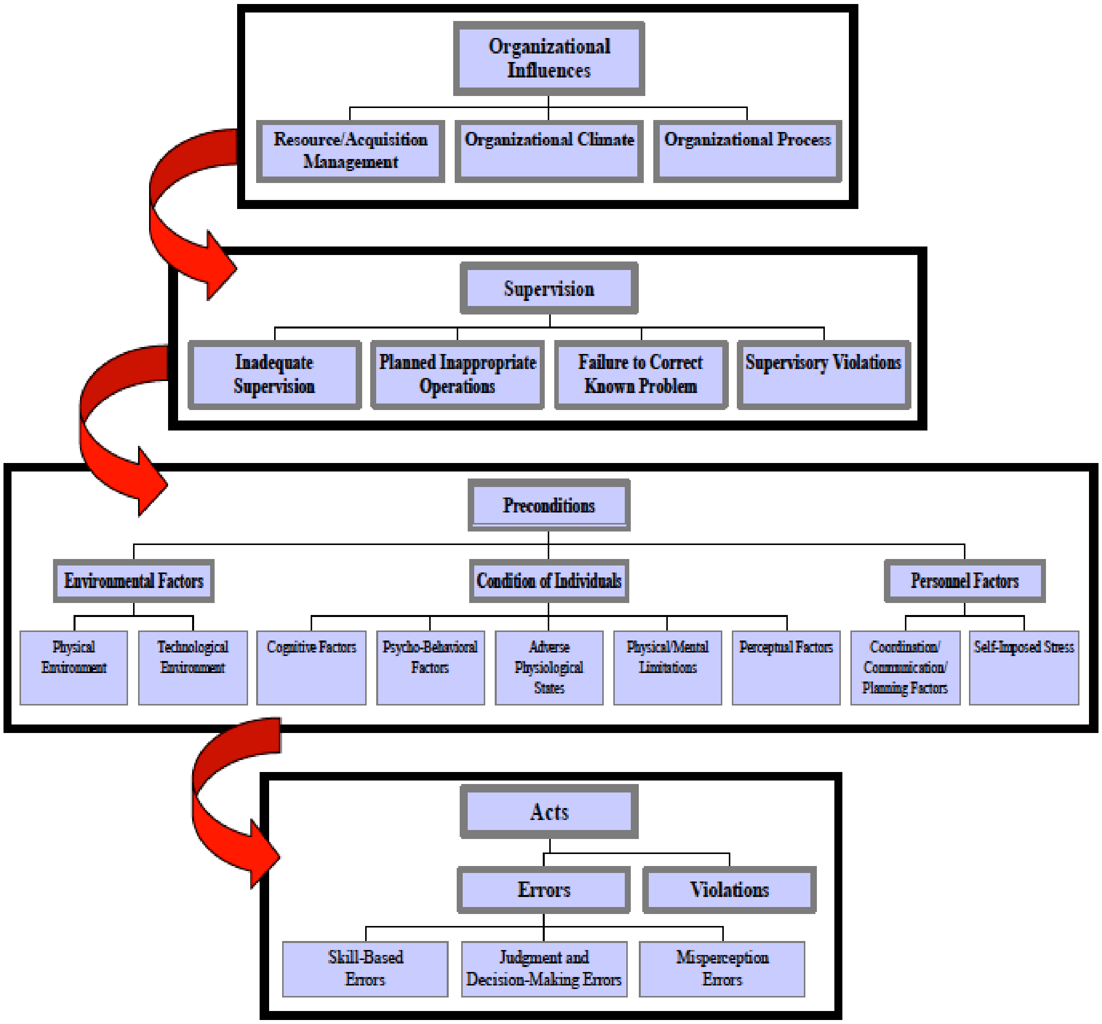

Figure 5. HFACS Taxonomy (From DoD, 2005)

1. Level I Acts

Acts are those factors that are most closely tied to the mishap, and can be described as active failures or actions committed by the operator that result in human error or unsafe condition (DoD, 2005). To assume that a mishap stems from a single point of failure is naïve (Woods, 2006). Reason (1990) suggests that mishaps are associated with numerous unsafe acts, too many to actually measure. The fact remains, however, that maritime mishaps and, maybe more importantly, near misses, are too often loaded with unsafe acts, which may be unintended (errors) or intended (violations) (Maurino et al., 1995).

The Acts level can be generalized into two categories, errors and violations, which are based on of the first level of defense in Reason's SCM. Errors are described in HFACS as the mental or physical errors that represent the mental or physical activities of individuals that fail to achieve their intended outcome (DoD, 2005). This is broken down for granularity into the subcategories of decision errors, judgment and decision-making errors, and misperception errors.

Skill-based errors represent sensory-motor performance during acts or activities that, following a statement of intention, take place without conscious control as smooth, automated, and highly integrated patterns of behavior (Rasmussen, 1983). This means actions are fundamentally accomplished without much conscious thought, but operators are then easily susceptible to forgetting steps or errors in techniques. A common mantra heard throughout the military is to "train like we fight and fight like we train." Meaning procedures followed in training should be identical to that in actual operations to avoid developing inappropriate habit patterns that could show up in combat. This is similar to "errors in techniques," since operators may have developed bad habits over time, which may lead to a failure later.

Judgment and decision-making errors can be viewed as procedural errors, poor choices, or problem-solving errors. Wiegman and Shappell described decision errors as "honest mistakes" or actions (or inactions) of individuals who had their "hearts in the right place" (Wiegman and Shappell, 2000, p.4) Misperception errors occur when an individual's perception of a situation differs from reality. Mariners often cite nighttime operations as being more difficult than daytime operations due to the loss of depth perception in the dark.

Violations are defined as the willful disregard for the rules and regulations (Rothblum et al., 2002). Violations are divided into two subcategories of routine and exceptional. Typical violations are common noncompliance to rules set forth for safety. DoD defines a violation as factors in a mishap when the actions of the operator represent willful disregard for rules and instructions, and lead to an unsafe situation (DoD, 2005). An example of a violation would be a ship transiting through a “no-wake zone” near a harbor, where there is a speed restriction of 10 knots. Ships may routinely ring up a couple of extra knots on the engines in order to get to their destination faster. Another violation may be a ship transiting at flank speed through the zone, with no regard for the safety of other vessels. While one violation may seem less harmful than the other, both are considered willful disregard for the rules, and therefore a violation.

2. Level II Preconditions for Unsafe Acts

DoD describes preconditions as factors in a mishap if latent and/or active preconditions affect practices, conditions, or actions of individuals and result in human error or an unsafe situation (DoD, 2005). Previous versions of HFACS included only two categories: substandard practices of operators and substandard conditions of operators. The DoD HFACS has three categories: environmental factors, condition of individuals, and personnel factors.

Environmental factors are assigned into two subcategories of technological or physical environments. The considerations of technological environment are issues such as the equipment and control designs; and human systems interface characteristics (displays), checklists, task factors and automation (Shappell et al., 2007). The physical environment incorporates issues such as weather in an operational setting, or the ambient environment such as temperature, vibration, and lighting (Shappell et al., 2007).

Much like the entire Preconditions level, the Condition of Individuals subcategory has gone through several iterations. Previous versions included only the subcategories of adverse mental states, adverse physiological states, and physical mental limitations (Shappell et al., 2007). The DoD HFACS has five separate subcategories: cognitive factors, psycho-behavioral factors, adverse physiological states, physical/mental limitations, and perceptual factors (DoD, 2005).

Cognitive factors are conditions affecting the perception or performance of an individual, resulting in human error (DoD, 2005). Inattention, confusion, and task oversaturation are some examples of cognitive factors. When an individual's personality traits, psycho-social problems, psychological disorders, or the wrong motivation causes an unsafe situation or a mishap, it is considered to be a psycho-behavioral factor (DoD, 2005). Examples of an adverse physiological state are factors such as fatigue, motion sickness, operating under the influence of prescription drugs, and preexisting illnesses or injuries. Physical/mental limitations are factors where the individual may not have the physical or mental capacity to handle a given situation, thereby causing a mishap. The final subcategory of perceptual factors considers the individual's misperception of surrounding operational conditions, visual illusions, or their disorientation in the operational environment.

The third category of Preconditions is broken into two subcategories of Coordination/Communication/Planning Factors and Self-Imposed Stress. Coordination/Communication/Planning factors looks at the interactions between the individuals and teams who are involved in the preparation and execution of an operation that resulted in human error (DoD, 2005). The subcategory of Self-Imposed Stress is when an operator is involved in an accident or unsafe situation due to their disregard of their lack of mental or physical readiness to perform. Physical-fitness level, alcohol, drugs, and diet are some of the considerations for this subcategory.

3. Level III Supervision

The next level of HFACS deals with supervision. This considers the role of inappropriate supervision to the extent that it causes an unsafe situation or a mishap. It is divided into four categories of inadequate supervision, planned inappropriate operations, failure to correct known problem, and supervisory violations.

A supervisor's role should be to provide subordinates with opportunities to succeed by providing guidance, training opportunities, and leadership as well as motivation (Shappell & Wiegmann, 2000). Navy officers typically receive BRM training, but this does not guarantee leadership performance. There are still issues with personality conflicts, policy, and supervision on ships, which can develop into a latent error, causing an unsafe situation.

In normal operations, a ship's crew ordered to conduct a mission that it is not trained to handle, is an unacceptable risk (DoD, 2005). High operational tempos and saturated schedules can jeopardize crew rest, which can subsequently affect the performance of the crew and create an unsafe situation. Planned inappropriate operations are also an issue when hazards are not addressed or are inadequately addressed by leadership.

The third category of known unsafe supervision is Failed to Correct a Known Problem. Supervisors may recognize issues with subordinates, material conditions of equipment, or training, which may affect the overall safety of the ship's system or crew (Shappell & Wiegmann, 2000). Even though the deficiencies are recognized, they may be ignored. Failing to correct the behavior or degradation in equipment can lead to errors. Supervisory violations are the willful disregard by supervisors of rules and regulations. These violations are considered rare and difficult to identify, but often instigate a chain of events for a mishap to follow (Shappell & Wiegmann, 2000).

4. Level IV Organizational Influences

Organizational Influences is the fourth level of the HFACS taxonomy. It is further divided into three categories: Resource Management, Organizational Climate, and Organizational Process. Organizational influences are latent conditions involving communication practices; the actions, omissions, or policies of upper-level management, which affect supervision; and the conditions or actions of the crew, resulting in a mishap (DoD, 2005).

Shappell and Wiegmann (2000) described Resource Management in the HFACS Final Report as a category that encompasses the realm of corporate-level decision making regarding the allocation and maintenance of organizational assets such as human resources (personnel), monetary assets, and equipment/facilities. The notion of this category is that operational resources should be managed with safety-minded and cost-effective decisions.

The military's resources are currently being subjected to potential sequestration cuts looming in January 2013. Deputy Secretary of Defense Ashton Carter voiced concerns that the military may be left hollow due to reduction of forces, ships, and equipment in order to meet the sequestration cutting measures of the DoD's budget (Carter, 2012). Resource Management looks at the high-level decisions made, and the trickle-down effects on equipment, crews, and operations that may cause latent human errors to occur.

Organizational Climate refers to the working atmosphere within the organization (Shappell & Wiegmann, 2000). This can also be defined as the organization's situational consistencies in the treatment of individuals. This refers to how an organization is structured, the policies in place regarding the attainment of sailors and officers, and the overall culture of the organization (DoD, 2005).

The final category in Level IV is Organizational Process. This looks at the overall operations (operational tempo, time pressure, schedules), procedures (standards and instructions) and oversight (Operational Risk Management [ORM] and safety programs), which can affect the safety of the ship and crew. High operational tempos and increased deployment lengths, may stress crews beyond their manning capabilities, subsequently causing conditions for latent errors to occur. Organizational processes can result in unrecognized hazards, leading to human error (DoD, 2005).

F. SUMMARY

The literature corroborates the need for continued research into maritime accidents. Human error is prevalent in commercial and Navy maritime accidents. Lin (1998) conveyed that previous papers expose human error causal factors as a significant cause of ship groundings, but few historical accident data records have the necessary information required for human factors analysis. This is partly due to the lack of effort to record information from ship groundings. Lacy (1998) found that the HFACS taxonomy supported over 92% of the causal factors found in the Naval Safety Center surface mishap reports, but recommended that there should be better HFACS causal factor descriptions to support analysis. In 2002 the HFWG revealed that human error contributes to up to 79% of tanker accidents, 89% to 96% of collisions, and 75% of allisions (Rothblum et al., 2002). Their research exposed that human factor issues abound in mishaps, and continued research is necessary to determine intervention strategies (Rothblum et al., 2002).

There are several definitions of human error. Rasmussen (2003) considered errors as the difference between an actual and a desired state. The authors in the book *Behind Human Error* suggest that the term “human error” hides too much about the system (Woods et al., 2010). Other research conveys the notion that inappropriate human behavior reduces system effectiveness and safety, which may result in an injury (Wickens et al., 2004). Reason (1990) suggests that there are too many definitions of human error, with no universal definition and no universal taxonomy to classify the term. Reason (1990) divided errors into two different categories of errors and violations. He considered violations as intentional human acts, which can be further divided into several categories such as routine, optimizing, and exceptional. In the same approach as violations, Reason (1990) divided errors into three basic types: knowledge-based mistakes; rule-based mistakes; and skill-based slips and lapses. These theories follow Rasmussen’s (1983) levels of performance of knowledge-based, rule-based, and skill-based behaviors.

Heinrich (1941) developed the Domino theory where accidents are the result of a sequence of events. The five dominos in his theory are lined up sequentially and if one is knocked over, a chain of events could follow, unless there is a subsequent removal of one of the unsafe acts or events (Heinrich, 1941). Reason’s SCM is a theory suggesting that the “holes in the cheese” are weaknesses in defensive barriers and if the holes were to line up, an accident may result (Reason, 1990). The holes in the cheese are due to some form of active failure or latent condition. Active failures have the most direct impact in accidents since it is due to the actions of operators. Latent conditions take time to develop and are difficult to identify since it represents hidden unsafe actions.

HFACS was originally developed for the aviation community as taxonomy to classify human errors. The backbone of HFACS is Reason's (1990) concept of active and latent failures. There are four levels of failure: Unsafe Acts (ACTS), Preconditions for Unsafe Acts (PRECONDITIONS), Unsafe Supervision (SUPERVISION), and Organizational Influences (ORGANIZATION). Acts are active failures most closely tied to the mishap (DoD, 2005). Preconditions are factors in a mishap if latent and/or active preconditions of individuals exist, resulting in human error. Supervision considers the supervisor's role in an unsafe situation or mishap (DoD, 2005). Organization is the fourth level of HFACS and is considered to be latent conditions involving upper-level management that affect supervision, conditions, or the actions of a crew (DoD, 2005).

III. METHOD

A. RESEARCH APPROACH

This research entails the *post hoc* analysis of the U.S. National Transportation Safety Board (NTSB), Canada's Transportation Safety Board (TSB), and the United Kingdom's (U.K.'s) Marine Accident Investigation Bureau (MAIB) marine accident database cases. Each reported accident case in the respective databases provides a summary of findings that indicate which factors were likely causes for an accident. The research approach employed the DoD HFACS to code human error causal factors for each reported marine accident case. Two raters independently coded each maritime accident case, and after coding was completed, inter-rater reliability was evaluated. Exploratory analysis was then conducted to identify significant human error patterns. Recognized patterns were further evaluated in terms of risk, for prioritizing intervention and suggesting control measure development.

B. DATABASE AND ACCIDENT CODING

1. Maritime Accidents

The maritime accident report data used was extracted from the U.S. NTSB, the Canadian TSB, and the U.K. MAIB maritime accident databases. The following paragraphs characterize each respective governmental safety agency, their investigation processes, and respective reporting procedures that generated the accident reports used.

Within the auspices of the U.S. Department of Transportation (DOT), the NTSB was developed as an independent administrative agency to promote safety among the different modes of transportation (National Transportation Safety Board [NTSB], 2012). Less than a decade later, the NTSB was separated from the DOT in order to eliminate any possible conflicts of interest, giving it the capability to conduct investigations and provide recommendations to the DOT (NTSB, 2012). The NTSB investigates mishaps involving all modes of transportation including aviation, highway, rail, marine, pipeline, and multimodal transportation (NTSB, 2012). The NTSB cooperates with the U.S. Coast Guard (USCG) through a Memorandum of Understanding to conduct investigations (NTSB, 2012).

The NTSB's investigation team is typically comprised of specialists in various disciplines. The "Go-Team," is the term used for the investigative team on standby to conduct an investigation 24 hours a day, 7 days a week (NTSB, 2012). In the event that a significant marine accident occurs, the team is dispatched to the scene where it can quickly assess the situation, provide any recommendations, and commence the investigation. An Investigator-in-Charge (IIC), who is a senior investigator from the NTSB, takes the lead of all the investigators who, in turn, head Individual Working Groups (NTSB, 2012). These groups are formed to divide the responsibilities of the investigation, which are based on areas of specialty including operations, structures, power plants, systems, weather, human performance, and survival. Human performance, weather, and survival factor specialists respond to all types of accidents (NTSB, 2012). Investigation lengths vary depending on the size of the accident, and final report delivery depends on the complexity of the analysis. Once the NTSB approves the report, it is added to its database for public viewing (NTSB, 2012).

The Canadian TSB conducts investigations on accidents related to marine, pipeline, rail, and air modes of transportation. The agency is an independent federal organization generated to improve transportation safety. TSB's mandate differs from organizations such as Transport Canada, the Canadian Coast Guard (CCG), and the Department of National Defense (DND), in that it makes recommendations in order to eliminate or reduce safety deficiencies in the various transportation modes (Government of Canada, 2012). Unlike the NTSB, which may work with the USCG in investigations, the TSB does not work with the CCG to investigate and to determine mishap causal factors.

The TSB investigation is broken down into three main phases: the Field phase, the Post-Field phase, and the Report Production phase (Government of

Canada, 2012). Once a decision to investigate an accident is made, an IIC is appointed and an investigation team consisting of operations, equipment, maintenance, engineering, scientific, and human performance specialists is assembled. The Field phase may last as little as a day or as long as several months. The Post-Field phase commences once the TSB departs the accident scene and may take several months, during which the team examines all of the collected information. Upon completing the Post-Field phase, the IIC produces a draft confidential report that is disseminated to interested parties for comment before the final report is prepared for eventual public release.

The U.K. has a distinct organization that conducts investigations of mishaps, the MAIB, which was established under the Merchant Shipping Act of 1995. The MAIB conducts investigations on all types of marine accidents within U.K. territorial waters and on registered U.K. vessels throughout the world. Their purpose is to investigate accidents to determine the cause and prevent future accidents. The MAIB does not apportion liability or blame; therefore, it does not enforce or prosecute under U.K. laws and regulations (MAIB, 2012). The MAIB operates under the umbrella of the Department of Transportation, but is not associated with the Maritime and Coastguard Agency (MCA). The head of the MAIB reports to the Secretary of State for Transport. MAIB's headquarters is located in Southampton and it has four investigation teams that are dispatched to worldwide locations. Each team is assigned a principle inspector and three inspectors, all of who are qualified in various disciplines of the marine industry.

The framework for MAIB's investigating and reporting of accidents are set forth in regulations by the U.K.'s Merchant Shipping Accident Reporting and Investigation provisions of 2005. Additionally, there is a Memorandum of Understanding among the MCA, MAIB, and the Health and Safety Executive when determining which organization will take the lead during investigations of common interest. If the MAIB decides to fully investigate an accident, its findings are publicly released in a MAIB accident report. It is not the intent of these reports to be used in court proceedings; rather, the MAIB aims to communicate to the public any safety issues and recommendations to prevent the same mishaps from occurring in the future (MAIB, 2012).

All accidents involving commercial vessels designed to carry passengers and cargo from the NTSB, TSB, and MAIB databases were screened. Reported navigational accidents where vessels that were underway collided with other vessels (collisions), collided with stationary objects (allisions), or ran aground were selected for analysis. The final database contained 48 navigational accidents between January 1, 2006 and December 21, 2011.

2. HFACS

The Joint Services Safety Chiefs (JSSC) directed that all services (through a Memorandum of Agreement between the services) use the DoD HFACS for their respective accident investigations (Joint Services Safety Chiefs [JSSC], 2012). The HFWG was charged with adapting the DoD HFACS document for joint use. The DoD HFACS is currently in use by the Naval Safety Center to investigate Naval aviation and maritime mishaps. Level I (Acts) is closely tied to the accident and is divided into two categories: errors and violations (DoD, 2005). Level II (Preconditions) considers environmental factors, the condition of individuals, and personnel factors (DoD, 2005). Level III (Supervision) considers accidents that can be traced back to the supervisory chain of command. It is divided into four subcategories of inadequate supervision, planned inappropriate operations, failure to correct known problem, and supervisory violations (DoD, 2005). Level IV (Organizational Influences) traces mishap causal factors to the fallible decisions of upper-level management, which affect supervisory practices as well as the conditions and actions of operators (DoD, 2005). This is further divided into three subcategories of resource/acquisition management, organizational climate, and organizational process. The DoD HFACS taxonomy provides subcategories within each category, with “nanocodes” to aid the investigator in classifying specific events.

3. HFACS Coding

Two raters independently evaluated the findings of each accident. The raters were selected based on their experience, knowledge, and background of the inner workings of surface ships in the U.S. Navy. Both raters are currently Naval officers, qualified as SWOs and studying the HSI curriculum at the NPS in Monterey, California. Collectively, they had approximately 25 years of experience and each had gained extensive knowledge of human error and HFACS through their studies of the HSI curriculum at NPS. The raters studied the NTSB accident cases and independently coded each case using HFACS. Upon completion of their review, the assessors' results were reviewed and a discussion took place to alleviate disagreements. After completing disagreement mediation, the amount of findings without HFACS classification was counted to determine the percentage of unclassified findings.

Each assessor was provided with the DoD HFACS guidance to review. Additional HFACS training was given to the raters to ensure that both raters understood the classification procedures and to clear up any preconceived notions regarding HFACS. Upon completion of training, each rater was given a copy of the NTSB maritime accident reports. Each rater worked independently to determine human error causal factors for each report. The codes used were in accordance with the HFACS instructions.

The DoD HFACS was adopted for use in this study for data analysis. In implementing the HFACS taxonomy, the two raters used the structure provided in the Naval Safety Center DoD HFACS flip book in order to facilitate the classification process. Table 1 illustrates the breakdown of the HFACS taxonomy into two levels. Level I is divided into Acts, Preconditions, Supervision, and Organizational Influences. Level II groups subcategories into each respective level.

LEVEL I		LEVEL II	
CATEGORIES	CODE	SUBCATEGORIES	CODE
Acts	A	Skill-Based Errors	AE1
		Judgment and Decision-Making Errors	AE2
		Perception Errors	AE3
		Violations	AV
Preconditions	P	Physical & Technological Environment	PE
		Coordination/Communication/Planning Factors & Self-Imposed Stress	PP
		Awareness (Cognitive) Factors, Psycho-Behavioral Factors, Adverse Physiological States, Physical/Mental Limitations, & Perceptual Factors	PC
Supervision	S	Inadequate Supervision	SI
		Failure to Correct Known Problem	SF
		Planned Inappropriate Operations	SP
		Supervisory Violations	SV
Organization	O	Resource/Acquisition Management	OR
		Organizational Climate	OC
		Organizational Process	OP
Not Applicable	N/A	Not Applicable	N/A

Table 1. DoD HFACS Analysis Grouping (After DoD, 2005)

C. DATA ANALYSIS

1. DOD HFACS Category Frequency

The frequency of occurrence for the DoD HFACS categories was assessed for each of the findings within the 48 maritime accident cases. The presence of a DoD HFACS nanocode was assessed as a one (1) and the absence of a nanocode will be assessed as a zero (0). No nanocode was used more than once in any accident case to avoid oversaturation of the data analysis. The codes were used to determine how often the DoD HFACS categories were used in each of the accidents. The total of the category assignments for all findings was calculated as a percentage to reflect the frequency of occurrence for the HFACS categories. HFACS categories with the highest frequency of occurrence were examined to identify patterns.

2. Inter-Rater Reliability

Cohen's Kappa analysis was conducted to determine inter-rater agreement of HFACS classifications (Curdy, 2009). From this, determinations can be made of where any disagreements between the two raters may be within the rating system. The reasons why there is disagreement and the frequency of disagreement between the two raters was examined and utilized as a dataset for further analysis. Cohen's Kappa was used to determine the proportion of agreement versus chance between two raters used for the HFACS coding of the maritime accidents. The Kappa coefficient is regarded as the choice statistical measurement for determining agreement between raters (Uebersax, 1987).

After the raters have classified the marine accidents using HFACS, kappa was computed to determine the level of agreement between the raters after determining the proportion of chance agreements. A kappa value of +1 shows 100% agreement between the two raters. A kappa value of 0 means there is not a relationship between the two raters, while a kappa of –1 is considered to be a 100% disagreement. Further interpretation of Cohen's Kappa values were broken down as follows: between 0.8 and 1 is considered Very Good, between 0.6 and 0.8 is considered Good, between 0.4 and 0.6 is considered Moderate Agreement, between 0.2 and 0.4 is considered Fair Agreement, and between 0 and 0.2 is considered Slight Agreement (Curdy, 2009).

3. Human Error Pattern Analysis

After the raters have conducted the HFACS coding on the NTSB maritime accident reports, a pattern analysis was conducted to determine what causal factors were identified most often. These results of the HFACS coding process from the raters were useful to categorize causal factors for further analysis in this research. A Chi Squared (χ^2) analysis was conducted to the examine the null hypothesis (H_o) of homogeneity that collisions, allisions, and groundings are equally likely to be categorized as findings of Acts (A), Preconditions (P), Supervision (S), Organization (O), or Not Applicable (N/A). Logistic regression analysis was then used to identify any significant patterns present in the coded accident data sets. Frequency tables were developed to represent how often an accident type has an HFACS Level I and Level II category associated with it.

Two rounds of logistic regression analysis were performed to determine if any HFACS Level I categories, or Level II subcategories, were significant for particular types of accidents. The first round consisted of testing specific accident types (first, collisions against allisions and groundings; then, allisions against collisions and groundings; and finally, groundings against collisions and allisions) and HFACS Level I categories. If a category was found to be statistically significant, further logistic regression was conducted based on the second level of HFACS. The second round of analysis looked at specific accident types, compared with HFACS Level II subcategories.

4. Risk Analysis

A risk assessment was conducted based on the percentage of agreements noted between the two raters and results of the follow-on analysis. This risk assessment was based on patterns and prevalent factors identified in the HFACS coding process. The risk assessment process detected common hazards noted in previous analysis and assessed any associated risks to vessels. Likelihood was considered as how often a specific hazard resulted in an accident, while severity considered the outcome of an accident if a hazard was the identified cause. For the purposes of this research, suggested accident probability and severity levels from the DoD Standard Practice for System Safety (DoD MIL-STD-882D, 2000) were used (see Tables 2 and 3).

Description*	Level	Specific Individual Item	Fleet or Inventory**
Frequent	A	Likely to occur often in the life of an item, with a probability of occurrence greater than 10^{-1} in that life.	Continuously experienced.
Probable	B	Will occur several times in the life of an item, with a probability of occurrence less than 10^{-1} but greater than 10^{-2} in that life.	Will occur frequently.
Occasional	C	Likely to occur some time in the life of an item, with a probability of occurrence less than 10^{-2} but greater than 10^{-3} in that life.	Will occur several times.
Remote	D	Unlikely but possible to occur in the life of an item, with a probability of occurrence less than 10^{-3} but greater than 10^{-6} in that life.	Unlikely, but can reasonably be expected to occur.
Improbable	E	So unlikely, it can be assumed occurrence may not be experienced, with a probability of occurrence less than 10^{-6} in that life.	Unlikely to occur, but possible.

Table 2. Accident Probability Criteria (From DoD MIL-STD-882D, 2000)

Description	Category	Environmental, Safety, and Health Result Criteria
Catastrophic	I	Could result in death, permanent total disability, loss exceeding $1M, or irreversible severe environmental damage that violates law or regulation.
Critical	II	Could result in permanent partial disability, injuries or occupational illness that may result in hospitalization of at least three personnel, loss exceeding $200K but less than $1M, or reversible environmental damage causing a violation of law or regulation.
Marginal	III	Could result in injury or occupational illness resulting in one or more lost work days(s), loss exceeding $10K but less than $200K, or mitigatible environmental damage without violation of law or regulation where restoration activities can be accomplished.
Negligible	IV	Could result in injury or illness not resulting in a lost work day, loss exceeding $2K but less than $10K, or minimal environmental damage not violating law or regulation.

Table 3. Accident Severity Criteria (From DoD MIL-STD-882D, 2000)

After a hazard was classified in terms of its probability and severity, it was assigned a Risk Assessment Code (RAC) taken from the Risk Assessment Matrix (see Table 4). These values were assigned to a risk category as either being HIGH, SERIOUS, MEDIUM, or LOW (see Table 5).

SEVERITY PROBABILITY	Catastrophic	Critical	Marginal	Negligible
Frequent	1	3	7	13
Probable	2	5	9	16
Occasional	4	6	11	18
Remote	8	10	14	19
Improbable	12	15	17	20

Table 4. Risk Assessment Matrix (From DoD MIL-STD-882D, 2000)

Mishap Risk Category	Mishap Risk Assessment Value
HIGH	1 - 5
SERIOUS	6 - 9
MEDIUM	10 - 17
LOW	18 - 20

Table 5. Mishap Risk Categories (After DoD MIL-STD-882D, 2000)

After significant factors were identified using logistic regression analysis, a modified hazards analysis was constructed. This provided a means to define what patterns were prevalent, determine the effects, and assess the risk associated with each latent condition. Patterns were evaluated based on the significant factors and associated accident reports. Risk was assessed in terms of probability and severity using the inputs provided by the two raters during the HFACS coding. The modified hazard analysis was conducted to increase awareness of potential safety issues and present potential corrective measures.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. RESULTS

A. ACCIDENT DATABASE

A total of 48 merchant ship accidents from three different countries were reviewed during the period of January 2006 through December 2011, and were selected for analysis based on them involving a collision, allision, or grounding. Of these 48 mishaps, 9 (12.5%) were from the U.S. NTSB database, 6 reports (18.8%) were from the Canadian TSB database, and 33 (68.8%) reports were from the U.K. MAIB database. Table 6 presents the distribution of the types of accidents pulled from each country's respective database.

	Number of Accidents	Groundings	Collisions	Allisions
Canada TSB	6	3	1	2
U.S. NTSB	9	1	3	5
U.K. MAIB	33	13	10	10
TOTAL	48	17	14	17

Table 6. Distribution of Maritime Accidents by Type

There were 518 total causal factors cited in 48 mishap cases. Table 7 presents the number of accident reports per country and the number of causal factors (findings) cited, with subsequent averages per accident report. Canada's TSB had the lowest number of reported accident cases, six, in which 32 causal factors were cited. It also had an overall average of 5.3 causal factors per report. This was much less than the U.S. NTSB average of 15.3 causal factors in nine reports. The U.K. MAIB had 33 accidents, with an overall average of 10.8 causal factors per report.

Country	Total Number of Accident Reports	Total Number of Causal Factors	Mean Causal Factors per Accident Report
Canada TSB	6	32	5.3
U.S. NTSB	9	138	15.3
U.K. MAIB	33	358	10.8
TOTAL	48	518	10.5

Table 7. Distribution of Number of Causal Factors by Country

Each of the three respective countries' accident reports were categorized into one of three types: Collisions, Allisions, and Groundings. Table 8 shows three different categories for each country, along with the respective minima, maxima, and averages of causal factors cited for each accident. The NTSB accident reports had the highest averages of causal factors cited in all three categories, while Canada's TSB had the lowest averages.

Country	Accident Type	Number of Accidents	Number of Cited Causal Factors	Number of Causal Factors Cited per Accident		
				Min	Max	Mean
Canada TSB	Collision	1	5	5	5	5.0
	Allisions	2	10	4	6	5.0
	Groundings	3	17	5	7	5.7
U.S. NTSB	Collision	3	48	14	19	16.0
	Allisions	5	64	9	30	14.8
	Groundings	1	16	16	16	16.0
U.K. MAIB	Collision	10	122	5	28	12.2
	Allisions	10	104	6	22	10.4
	Groundings	13	132	3	24	13.2
TOTAL		48	518			

Table 8. Distribution of Number of Causal Factors by Accident Type

B. RATER CAUSAL FACTOR CODING

The raters were instructed to only use one HFACS nanocode for each of the accident reports' findings, in order to avoid oversaturation of the data analysis. The raters' nanocodes were then split into Level I and Level II codes for analysis. If the raters had findings that did not relate to the HFACS taxonomy, then the findings were not classified into the HFACS taxonomy and were annotated as Not Applicable (N/A).

The frequency of occurrence for the HFACS categories was assessed for each of the 48 maritime accident cases. There were a total of 18 Level I and II HFACS category assignments for all findings, which were considered causal factors. The percentage listed in Table 9 reflects the ratio frequency occurrence of a particular HFACS category to the total number of accidents (48) in Level I. The category with the highest frequency of occurrence in HFACS Level One/Rater 1 is the *Preconditions* category, with a 27% of occurrence. Rater 2 classified findings into *Organization* (28%) as the most common factor.

HFACS LEVEL I CATEGORY	RATER 1		RATER 2	
	#	%	#	%
ACTS	137	26.4	132	25.5
PRECONDITIONS	139	26.8	97	18.7
SUPERVISION	40	7.7	55	10.6
ORGANIZATION	134	25.9	146	28.2
NOT APPLICABLE	68	13.1	88	17.0

Table 9. Level I HFACS Distribution

Raters then coded the findings into the HFACS Level II subcategories, which was more difficult to conduct due to the increased number of possible subcategories. Table 10 reflects the frequency of occurrence for all HFACS categories and respective subcategories for Level II. In Level II, Rater 1 classified findings into the *Coordination/Communication/Planning Factors (PP)* subcategory (17.4%) in the HFACS Level I *Preconditions* category most often, while Rater 2 used the subcategory OP (27.8%) in the *Organization* category most often.

HFACS	LEVEL 2 SUBCATEGORIES	RATER 1		RATER 2	
		#	%	#	%
ACTS	AE1	37	7.1	58	11.2
	AE2	71	13.7	61	11.8
	AE3	5	1.0	2	0.4
	AV	24	4.6	11	2.1
PRECONDITIONS	PE	11	2.1	4	0.8
	PP	90	17.4	68	13.1
	PC	38	7.3	25	4.8
SUPERVISION	SI	18	3.5	27	5.2
	SF	1	0.2	5	1.0
	SP	14	2.7	16	3.1
	SV	7	1.4	7	1.4
ORGANIZATION	OR	38	7.3	28	5.4
	OC	10	1.9	11	2.1
	OP	86	16.6	107	20.7
NOT APPLICABLE	N/A	68	13.1	88	17.0

Table 10. HFACS Level II Distribution by Raters

C. INTER-RATER RELIABILITY

After the two raters' independent coding process was conducted, there were initially 226 (43.6%) agreements of the 518 findings for Level I and 234 (45%) for Level II. After discussion, mediation, and reconciliation of all disagreements, consensus was reached and agreement between raters was obtained for 405 (78%) findings in Level I and 357 (69%) in Level II. This was used as the dataset for further analysis. The remaining unresolved causal factors were considered Not Applicable (N/A) within the current HFACS taxonomy. The overall Cohen's Kappa values for each HFACS levels are given in Table 11. The best agreement between the two raters was in Level I (Cohen's Kappa = 0.72), where the findings were categorized as an Act, Precondition, Supervision, Organization, or Not Applicable (N/A). Level II agreement was lower (Cohen's Kappa = 0.64). This level represented 13 different categories and a Not Applicable category. The Cohen's Kappa Level I finding of 0.72 and Level II of 0.64 are both considered good agreement between the two raters.

HFACS Level	Before or After Discussion	Cohen's Kappa Value
I	Before	0.45
	After	0.72
II	Before	0.39
	After	0.64

Table 11. Cohen's Kappa Results

D. HUMAN ERROR PATTERN ANALYSIS

1. HFACS Level I Analysis

Figure 6 shows Rater 1 classified collisions as an HFACS category Acts more often than any other category (33.7%). Groundings classified into the *Supervision* category the least (6.1%). Null Hypothesis (H_o): The type of accidents (groundings, collisions, and allisions) is equally likely to be categorized as findings of Acts (A), Preconditions (P), Supervision (S), Organization (O), or Not Applicable (N/A). Using Rater 1 data in HFACS Level I, the null hypothesis of homogeneity of allision, collision, and grounding incidents, with regard to the distribution of HFACS categorical findings, is rejected (p = 0.03).

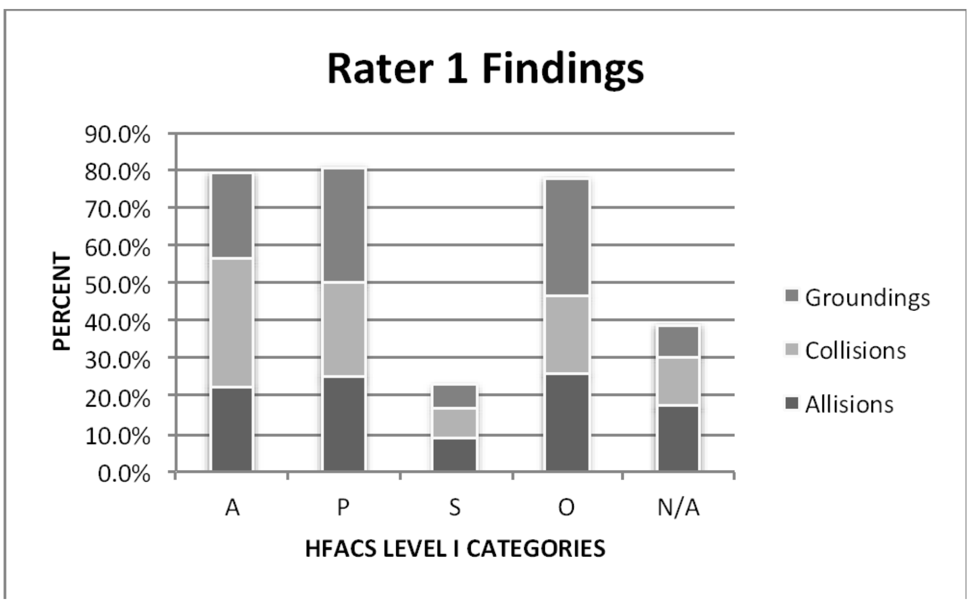

Figure 6. Percentage of Rater 1-Level I Findings by Accident Types

Figure 7 shows that Rater 2 classified groundings (33.3%) and allisions (28.1%) most often into the Organization category and collisions (34.3%) into the Acts category. Once again the Null Hypothesis (H_o): The type of accidents (collisions, allisions, and groundings) is equally likely to be categorized as findings of Acts (A), Preconditions (P), Supervision (S), Organization (O), or Not Applicable (N/A). Using Rater 2 data only, the null hypothesis of homogeneity of the type of accidents (groundings, collisions, and allisions), with regard to the distribution of HFACS categorical findings, is rejected (p = 0.01).

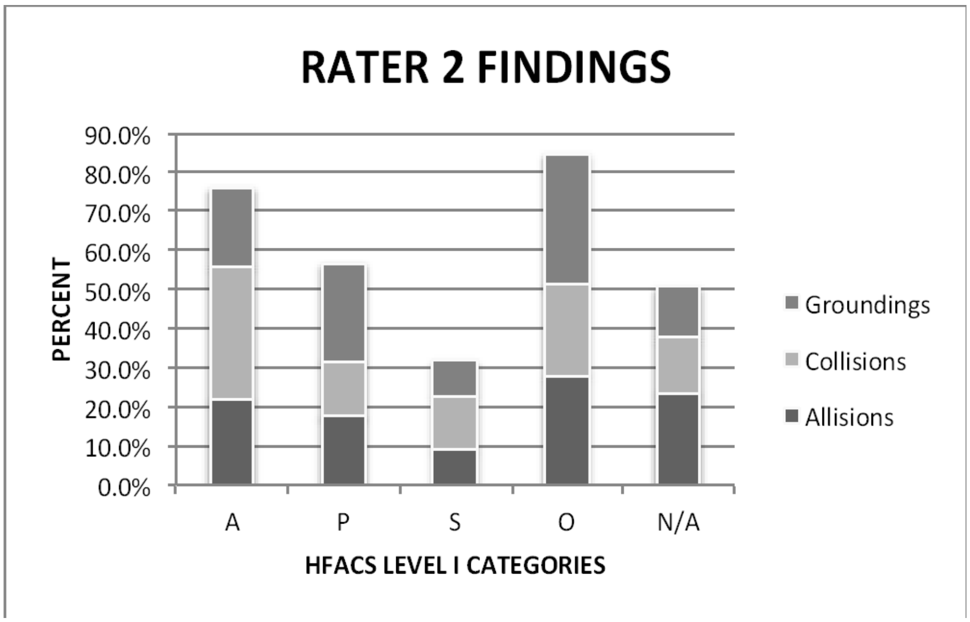

Figure 7. Percentage of Rater 2 Level I Findings by Accident Types

Figure 8 shows the percentage of accident report findings where the two raters agreed on the HFACS Level I categorization, with regards to the type of accident. Collisions were classified as Acts by both raters most often (36.1%) and groundings were classified as Supervision the least (6.1%).

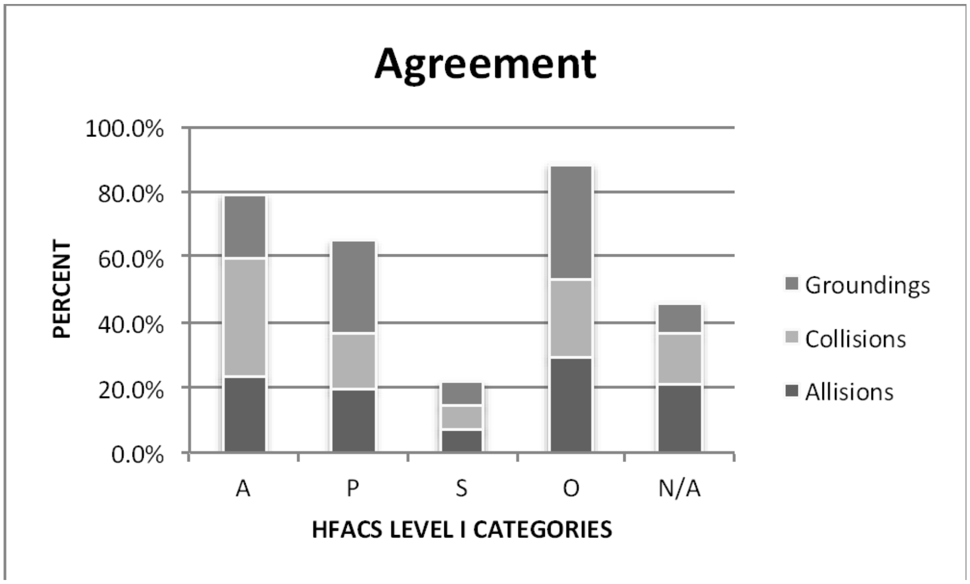

Figure 8. Agreement Between Raters by Accident Type in HFACS Level I

Table 12 shows how many mishaps were coded with at least one finding from the categories of HFACS Level I codes for each subsequent accident type. The raters used the HFACS Level I category of Organization most often among the Collisions (92.9%), Allisions (100%), and Groundings (100%). The Supervision category was used the least for Allisions (41.2%) and Groundings (41.2%).

HFACS Level I Categories								
Type of Accident	Organization		Supervision		Preconditions		Acts	
	#	%	#	%	#	%	#	%
Collision (n = 14)	13	92.9%	12	85.7%	12	85.7%	12	85.7%
Allision (n = 17)	17	100.0%	7	41.2%	15	88.2%	15	88.2%
Grounding (n = 17)	17	100.0%	7	41.2%	15	88.2%	15	88.2%

Table 12. Frequency of HFACS Level I Codes for Each Accident Type

2. HFACS Level II Analysis

Figure 9 represents the percentage of findings by Rater 1 classified in the HFACS Level II subcategories, with regards to the type of accident. Four factors were more frequent than the rest of the factors. The subcategory of Self-Imposed Stress (PP) represented the highest overall percentage when Rater 1 classified findings. Subcategory PP represented the highest usage by Rater 1 for allisions and groundings. Organizational Processes (OP) and Judgment and Decision Errors (AE2) were the two other factors that were prevalent in the figure. Not Applicable (N/A) was used by Rater 1 when unable to classify accident findings. This occurred for allisions (17.1%) most often when compared to collisions (13.1%) and groundings (8.5%).

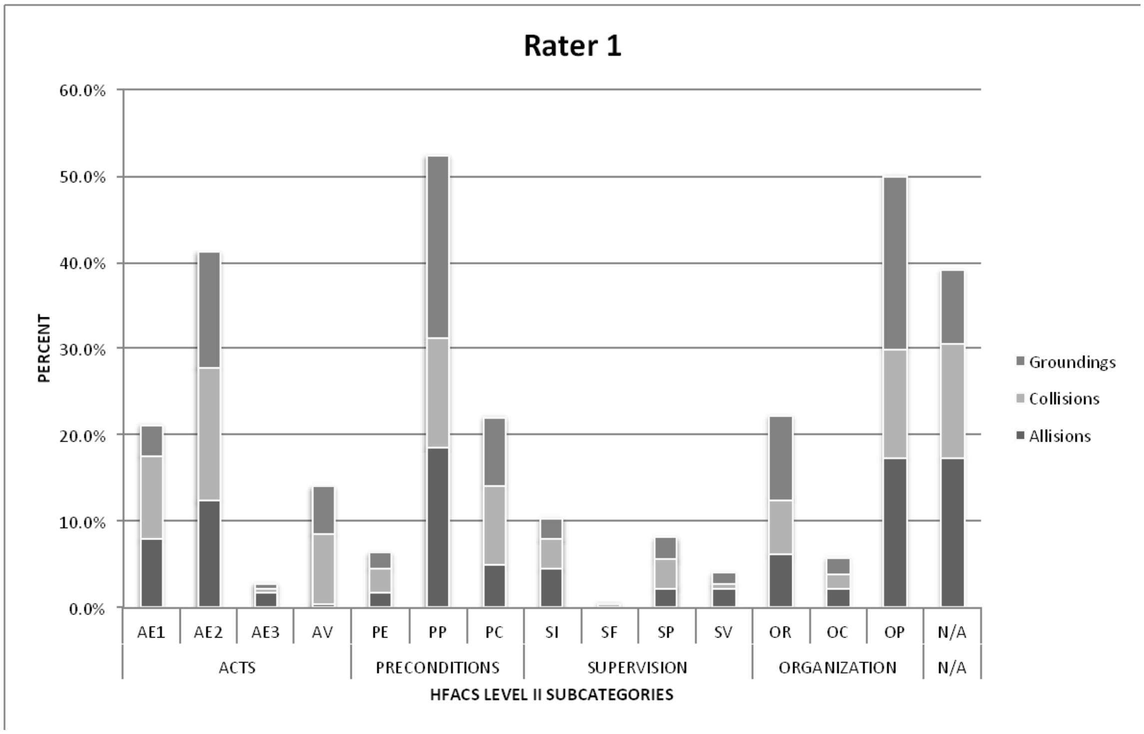

Figure 9. Percentage of Rater 1 Level II Findings by Accident Types

Figure 10 represents Rater 2's percentage of findings classified in the HFACS Level II subcategories, with regards to the type of accident. Not Applicable (N/A) was used by Rater 2 when unable to classify accident findings. This occurred for allisions (23.6%) most often when compared to collisions (14.3%) and groundings (12.7%). The subcategory of Organizational Processes (OP) had the highest percentage overall, and subsequently was rated the highest for collisions (16.0%) and groundings (24.2%). There were three instances where a subcategory was not used. The most significant differences noted between the raters are the use of subcategories of Skill-Based Errors (AE1), Judgment and Decision-Making Errors (AE2), and Coordination/Communication/Planning Factors (PP).

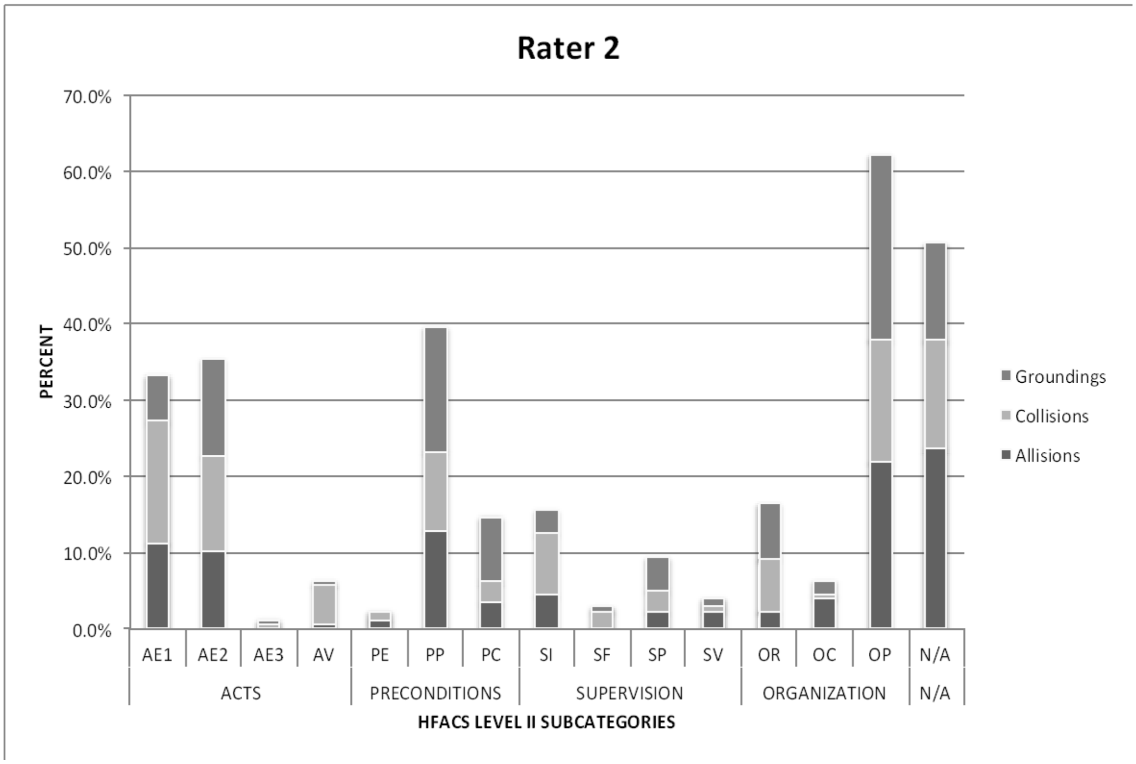

Figure 10. Percentage of Rater 2 Level II Findings by Accident Types

Figure 11 shows the percentage of accident report findings where the two raters agreed on the HFACS Level II categorization, with regards to the type of accident. Collisions classified into subcategory Organizational Processes (OP) resulted in the highest percentage of agreement (27.7%). There were nine instances where a subcategory was not used. Both raters did not classify allisions into the HFACS subcategory of Failure to Correct Known Problem (SF). Not Applicable (N/A) findings represented the highest agreement percentage for findings involving allisions (24.4% compared to all subcategories).

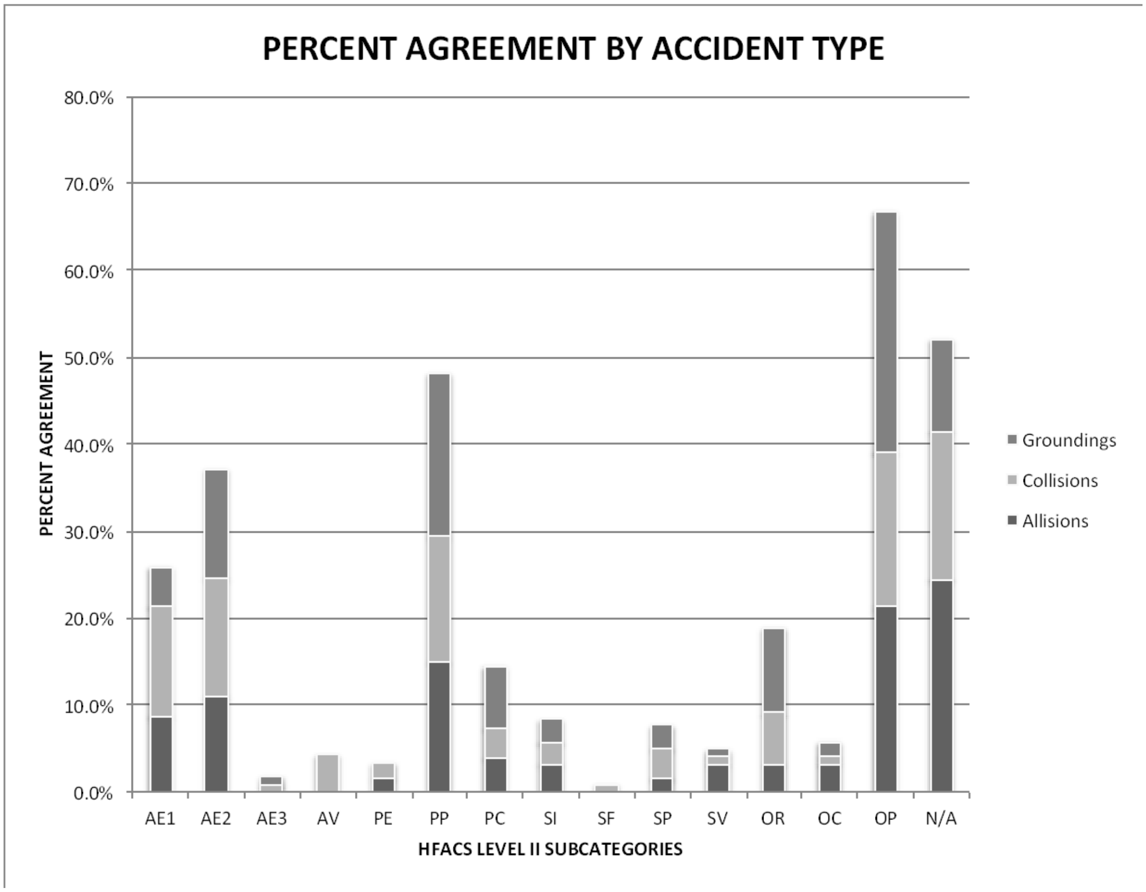

Figure 11. Agreement Between Raters by Accident Type in HFACS Level II

Table 13 represents how many mishaps were coded with at least one finding from the subcategories of HFACS Level II codes for each subsequent accident type. The raters used HFACS Level II subcategories of Perception Errors (AE3), Technological Environment (PE), Failure to Correct Known Problem (SF), and Supervisory Violations (SV) most often among the Collisions (93% each). For allisions and groundings, the Self-Imposed Stress (PP) was used the most often (82% each).

Type of Accident	OP		OC		OR		SV		SP		SF		SI		PC		PP		PE		AV		AE3		AE2		AE1	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Collision (n = 14)	2	14	12	86	7	50	13	93	10	71	13	93	8	57	9	64	6	43	13	93	9	64	13	93	5	36	5	36
Allision (n = 17)	10	59	2	12	6	35	1	6	4	24	0	0	7	41	8	47	14	82	2	12	2	12	2	12	10	59	10	59
Grounding (n = 17)	13	76	4	24	11	65	2	12	5	29	1	6	3	18	7	41	14	82	3	18	6	35	1	6	14	82	9	53

Note: # is Number of Accidents; % is associated percentages.

Table 13. Frequency of HFACS Level II Codes for Each Accident Type

3. Logistic Regression Analysis

The first round of logistic regression analysis was performed to determine if there was a significant difference in the HFACS Level I categories among the three types of accidents. Three areas were examined in the logistic regression model analysis output:

The first area examined was the whole-model test:

If $p < 0.05$, then the logistic model is NOT useful.

If $p > 0.05$, then the logistic model is useful.

The second area examined was the lack-of-fit test:

If $p > 0.05$, then the logistic model is adequate.

If $p < 0.05$, then the logistic is inadequate.

The final areas examined were the parameter estimates:

If $p < 0.05$, then the indicator is statistically significant.

If $p > 0.05$, then the indicator is NOT statistically significant.

(Note: All of the logistic regression tables were extracted from the JMP statistic software program output for each respective logistic regression model).

In testing collisions versus non-collisions across HFACS Level I categories (see Table 14), the logistic regression analysis revealed that the whole model test was slightly above being statistically significant ($\chi^2_{(4)} = 8.21$, $p = 0.08$), but the lack-of-fit test revealed that there was no evidence of a lack-of-fit ($p = 0.26$). In testing the significance of the predictors in the logistic model, there was a significant difference for the Supervision category ($p = 0.02$).

Nominal Logistic Fit for Collisions
Converged in Gradient, 5 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4.106587	4	8.213174	0.0841
Full	24.868001			
Reduced	28.974588			
RSquare (U)	0.1417			
AICc	61.1646			
BIC	69.092			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	5	3.255661	6.511322
Saturated	9	21.612340	Prob>ChiSq
Fitted	4	24.868001	0.2596

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	1.40090493	0.7524175	3.47	0.0626
A[0]	-0.5491632	0.5667807	0.94	0.3326
P[0]	-0.3582362	0.529224	0.46	0.4985
S[0]	1.18368231	0.5104219	5.38	0.0204
O[0]	0.85348375	0.6119582	1.95	0.1631

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
A	1	1	0.93604623	0.3333
P	1	1	0.44788038	0.5033
S	1	1	7.56688545	0.0059
O	1	1	2.45054296	0.1175

Table 14. Collisions Compared to All HFACS Level I Categories

The Supervision category was then singled out as the only independent variable and logistic regression was again conducted (see Table 15). The whole-model test reveals there is evidence to suggest that the model is useful ($\chi^2_{(1)}$ = 5.01, p = 0.03). In testing the significance of the Supervision category as a predictor in the parameter estimates, the category was determined to be statistically significant (p = 0.05).

Nominal Logistic Fit for Collisions
Converged in Gradient, 5 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.505260	1	5.010521	0.0252
Full	26.469328			
Reduced	28.974588			
RSquare (U)	0.0865			
AICc	57.2053			
BIC	60.6811			
Observations (or Sum Wgts)		48		

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	1.24245332	0.4187448	8.80	0.0030
S[0]	0.83698822	0.4187448	4.00	0.0456

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
S	1	1	5.01052067	0.0252

Table 15. Collisions Compared to HFACS Level I Supervision Category

Further the logistic regression was performed to determine if there was a significant difference between collisions and non-collisions across the four HFACS Level II subcategories of Supervision (see Table 16). The whole-model test provided evidence that the model was not useful for Supervision subcategories to differentiate collisions from non-collisions ($\chi^2_{(4)} = 6.85$, p = 0.14). Since the lack-of-fit test just passed (p = 0.06), the parameter estimates were examined. The parameter estimates revealed there was not a significant difference between collisions and noncollisions across the subcategories of Supervision.

Nominal Logistic Fit for Collision
Converged in Gradient, 4 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.425812	4	6.851624	0.1439
Full	25.548776			
Reduced	28.974588			
RSquare (U)	0.1182			
AICc	62.5261			
BIC	70.4536			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	4	4.429580	8.85916
Saturated	8	21.119196	Prob>ChiSq
Fitted	4	25.548776	0.0647

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.0052931	0.771106	0.00	0.9945
SI[0]	0.49541812	0.3709685	1.78	0.1817
SF[0]	1.05589804	0.6973629	2.29	0.1300
SP[0]	0.35934485	0.3779477	0.90	0.3417
SV[0]	-0.2151355	0.7157139	0.09	0.7637

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
SI	1	1	1.81146069	0.1783
SF	1	1	2.83295263	0.0923
SP	1	1	0.89971149	0.3429
SV	1	1	0.09496503	0.7580

Table 16. Collisions Compared to All HFACS Level II Supervision Subcategories

In testing the next accident type, allisions versus non-allisions across the HFACS Level I categories, the logistic regression revealed that the whole-model test was not statistically significant ($\chi^2_{(4)} = 7.25$, $p = 0.12$), but the lack-of-fit test reveals that there was no evidence of a lack of fit ($p = 0.51$). In testing the significance of the predictors in the logistic model, there was a significant difference for the Organization category ($p = 0.02$). Table 17 displays the logistic regression output for allisions versus nonallisions across all HFACS Level I categories.

Nominal Logistic Fit for Allisions
Converged in Gradient, 4 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.626992	4	7.253984	0.1231
Full	27.572426			
Reduced	31.199418			
RSquare (U)	0.1163			
AICc	66.5734			
BIC	74.5009			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	5	2.155063	4.310125
Saturated	9	25.417364	Prob>ChiSq
Fitted	4	27.572426	0.5057

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.4092816	0.740479	0.31	0.5805
A[0]	-0.1123596	0.4724423	0.06	0.8120
P[0]	0.17064028	0.5176082	0.11	0.7416
S[0]	-0.1897399	0.3654449	0.27	0.6036
O[0]	-1.3602422	0.603014	5.09	0.0241

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
A	1	1	0.0561238	0.8127
P	1	1	0.11208582	0.7378
S	1	1	0.26849351	0.6043
O	1	1	6.90667652	0.0086

Table 17. Allisions Compared to All HFACS Level I Categories

The Organization category was then singled out as the only independent variable and logistic regression was conducted again (see Table 18). The whole-model test reveals there was evidence to suggest the model was useful ($\chi^2_{(1)}$ = 6.74, p = 0.01). In testing the significance of the Organization category as a predictor in the parameter estimates, the category was determined to be statistically significant (p = 0.03).

Nominal Logistic Fit for Allisions
Converged in Gradient, 4 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.368728	1	6.737457	0.0094
Full	27.830690			
Reduced	31.199418			
RSquare (U)	0.1080			
AICc	59.928			
BIC	63.4038			
Observations (or Sum Wgts)		48		

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.3465736	0.5737305	0.36	0.5458
O[0]	-1.2628643	0.5737305	4.85	0.0277

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
O	1	1	6.73745672	0.0094

Table 18. Allisions Compared to HFACS Level I Organization category

Further regression analysis was performed to determine if there was a significant difference between allisions versus non-allisions and the three HFACS Level II subcategories of Organization (see Table 19). The whole-model test provided evidence that the model was useful for Organization subcategories to predict allisions ($\chi^2_{(3)}$ = 7.72, p = 0.05). Since the lack-of-fit test passed (p = 0.37), the parameter estimates were examined. The parameter estimates revealed that there was one subcategory, Resource/Acquisition Management (OR), as statistically significant (p = 0.03) between allisions versus nonallisions.

Nominal Logistic Fit for Allision
Converged in Gradient, 4 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	3.861150	3	7.722301	0.0521
Full	27.338268			
Reduced	31.199418			
RSquare (U)	0.1238			
AICc	63.6068			
BIC	70.1613			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	3	1.557820	3.115639
Saturated	6	25.780448	Prob>ChiSq
Fitted	3	27.338268	0.3741

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.53591277	0.4964264	1.17	0.2803
OR[0]	-0.714978	0.3331873	4.60	0.0319
OC[0]	-0.2683262	0.4648663	0.33	0.5638
OP[0]	-0.5431056	0.3639789	2.23	0.1357

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
OR	1	1	4.92916201	0.0264
OC	1	1	0.34971591	0.5543
OP	1	1	2.26521514	0.1323

Table 19. Allisions Compared to All HFACS Level II Organization Subcategories

The Resource/Acquisition Management (OR) subcategory was then singled out as the only independent variable and logistic regression was conducted again (see Table 20). The whole-model test reveals there was evidence to suggest the model was useful ($\chi^2_{(1)} = 4.73$, $p = 0.03$). In testing the significance of the Resource/Acquisition (OR) subcategory as a predictor in the parameter estimates, the subcategory was determined to be statistically significant ($p = 0.03$).

Nominal Logistic Fit for Allision
Converged in Gradient, 4 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.365079	1	4.730157	0.0296
Full	28.834340			
Reduced	31.199418			
RSquare (U)	0.0758			
AICc	61.9353			
BIC	65.4111			
Observations (or Sum Wgts)		48		

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.57872639	0.3182746	3.31	0.0690
OR[0]	-0.6740366	0.3182746	4.49	0.0342

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
OR	1	1	4.7301574	0.0296

Table 20. Allisions Compared to HFACS Level II Resource/Acquisition Management Subcategory

In testing the third accident type, groundings versus nongroundings across HFACS Level I categories, the analysis revealed that the whole-model test was statistically significant ($\chi^2_{(4)}$ = 9.82, p = 0.04). The lack-of-fit test revealed that there was no evidence of a lack of fit (p = 0.20). In testing the significance of the predictors in the logistic model, there was a significant difference for the Supervision category (p = 0.05). Table 21 displays the logistic regression output for groundings versus nongroundings across all HFACS Level I categories.

Nominal Logistic Fit for Groundings
Converged in Gradient, 16 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	4.911282	4	9.822564	0.0435
Full	26.288136			
Reduced	31.199418			
RSquare (U)	0.1574			
AICc	64.0048			
BIC	71.9323			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	5	3.643361	7.286723
Saturated	9	22.644775	Prob>ChiSq
Fitted	4	26.288136	0.2002

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
InterceptUnstable	8.86706627	1097.9158	0.00	0.9936
A[0]	0.51221673	0.5111431	1.00	0.3163
P[0]	0.11615528	0.5155746	0.05	0.8218
S[0]	-0.692883	0.3591672	3.72	0.0537
O[0]Unstable	8.09638134	1097.9157	0.00	0.9941

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
A	1	1	1.06563523	0.3019
P	1	1	0.05130339	0.8208
S	1	1	3.96381224	0.0465
O	1	1	3.30747541	0.0690

Table 21. Groundings Compared to All HFACS Level I Categories

The Supervision (S) category was then singled out as the only independent variable and regression analysis was conducted again (see Table 22). The whole-model test revealed there was evidence to suggest that the model was useful ($\chi^2_{(1)}$ = 5.07, p = 0.02). In testing the significance of the Supervision (S) category as a predictor in the parameter estimates, the category was determined to be statistically significant (p = 0.03).

Nominal Logistic Fit for Groundings
Converged in Gradient, 4 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.535927	1	5.071853	0.0243
Full	28.663492			
Reduced	31.199418			

RSquare (U)	0.0813
AICc	61.5937
BIC	65.0694
Observations (or Sum Wgts)	48

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.48322026	0.3206772	2.27	0.1318
S[0]	-0.7063638	0.3206772	4.85	0.0276

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
S	1	1	5.07185303	0.0243

Table 22. Groundings Compared to HFACS Level I Supervision Category

Further logistic regression was performed to determine if there was a significant difference between groundings and the three HFACS Level II subcategories of Supervision (see Table 23). The whole-model test provided evidence that the model was not useful for Supervision subcategories to predict groundings ($\chi^2_{(4)} = 4.87$, $p = 0.31$). Additionally, the lack-of-fit test did not pass ($p = 0.03$); therefore, the model was considered inadequate. Since the Inadequate Supervision (SI) subcategory was slightly above significance, it was singled out for further logistic regression analysis.

Nominal Logistic Fit for Groundings
Converged in Gradient, 4 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.403583	4	4.807166	0.3077
Full	28.795836			
Reduced	31.199418			
RSquare (U)	0.0770			
AICc	69.0202			
BIC	76.9477			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	4	5.193977	10.38795
Saturated	8	23.601858	Prob>ChiSq
Fitted	4	28.795836	0.0344

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.83809973	0.7304576	1.32	0.2512
SI[0]	-0.7337685	0.3899204	3.54	0.0599
SF[0]	-0.0134363	0.6913178	0.00	0.9845
SP[0]	-0.0885687	0.3531682	0.06	0.8020
SV[0]	0.09761493	0.5707673	0.03	0.8642

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
SI	1	1	4.05096864	0.0441
SF	1	1	0.00037875	0.9845
SP	1	1	0.06336507	0.8013
SV	1	1	0.02924941	0.8642

Table 23. Groundings Compared to All HFACS Level II Supervision Subcategories

The Inadequate Supervision (SI) subcategory was singled out as the only independent variable and regression analysis was conducted again (see Table 24). The whole-model test revealed that there was evidence to suggest the model was useful ($\chi^2_{(1)}$ = 4.73, p = 0.03). In testing the significance of the Inadequate Supervision (SI) category as a predictor in the parameter estimates, the category was determined to be statistically significant (p = 0.03).

Nominal Logistic Fit for Groundings
Converged in Gradient, 4 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	2.361617	1	4.723235	0.0298
Full	28.837801			
Reduced	31.199418			
RSquare (U)	0.0757			
AICc	61.9423			
BIC	65.418			
Observations (or Sum Wgts)		48		

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.87148465	0.3653521	5.69	0.0171
SI[0]	-0.7379533	0.3653521	4.08	0.0434

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
SI	1	1	4.72323459	0.0298

Table 24. Groundings Compared to HFACS Level II Inadequate Supervision Subcategory

The second round of logistic regression was performed to determine if there was a significant difference in the HFACS Level II subcategories among the three types of accidents. In testing collisions against HFACS Level II subcategories, the analysis revealed that the whole-model test was useful (p = 0.0001), but the lack-of-fit test revealed that there was no evidence of a lack of fit (p = 1.00). This proved the model was unstable, and therefore disregarded (see Appendix).

In testing allisions against HFACS Level II subcategories, the analysis revealed that the whole-model test was useful ($\chi^2_{(14)}$ = 26.02, p = 0.03). The lack-of-fit test revealed that there was no evidence of a lack of fit (p = 0.48). In testing the significance of the predictors in the logistic model, there was a significant difference for the Violations (AV) subcategory (p = 0.02). Organizational Processes (OP) was slightly above significance (p = 0.06); therefore, it was considered for further analysis. Table 25 displays the logistic regression output for Allisions compared to all HFACS Level II subcategories.

Nominal Logistic Fit for Allisions
Converged in Gradient, 17 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	13.013509	14	26.02702	0.0257
Full	18.185910			
Reduced	31.199418			
RSquare (U)	0.4171			
AICc	81.3718			
BIC	94.4398			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	31	15.413321	30.82664
Saturated	45	2.772589	Prob>ChiSq
Fitted	14	18.185910	0.4750

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
InterceptUnstable	8.40218346	1819.8362	0.00	0.9963
AE1[0]	0.14560446	0.5483299	0.07	0.7906
AE2[0]	-0.3140341	0.6618363	0.23	0.6352
AE3[0]Unstable	9.26644452	1822.3587	0.00	0.9959
AV[0]	-1.342555	0.7380337	3.31	0.0689
PE[0]	-1.1294389	0.8008807	1.99	0.1585
PP[0]	0.1408082	0.7250278	0.04	0.8460
PC[0]	-0.1850974	0.600532	0.10	0.7579
SI[0]	0.61969716	0.5178229	1.43	0.2314
SF[0]Unstable	-16.37111	2575.4212	0.00	0.9949
SP[0]	0.03877039	0.6103909	0.00	0.9494
SV[0]	0.73333437	1.255641	0.34	0.5592
OR[0]	-0.7786933	0.5199583	2.24	0.1342
OC[0]	-0.723262	0.664032	1.19	0.2761
OP[0]	-1.2732524	0.6857943	3.45	0.0634

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AE1	1	1	0.07112286	0.7897
AE2	1	1	0.22637875	0.6342
AE3	1	1	4.76113747	0.0291
AV	1	1	4.42253528	0.0355
PE	1	1	2.39774584	0.1215
PP	1	1	0.03798728	0.8455
PC	1	1	0.09689233	0.7556
SI	1	1	1.49570345	0.2213
SF	1	1	2.16529422	0.1412
SP	1	1	0.00404045	0.9493
SV	1	1	0.33500886	0.5627
OR	1	1	2.4341761	0.1187
OC	1	1	1.26760284	0.2602
OP	1	1	4.52760305	0.0334

Table 25. Allisions Compared to All HFACS LEVEL II Subcategories

The Violations (AV) and Organizational Processes (OP) subcategories were singled out as the only independent variables and logistic regression was conducted again (see Table 26). The whole-model test revealed that there was evidence to suggest the model was useful ($\chi^2_{(2)} = 11.53$, $p = 0.003$). The lack-of-fit test revealed that the model was adequate ($p = 0.07$). In testing the significance of the Violations (AV) subcategory as a predictor in the parameter estimates, the subcategory was determined to be statistically significant ($p = 0.01$). The subcategory of Organizational Processes (OP) was not considered statistically significant.

Nominal Logistic Fit for Allisions
Converged in Gradient, 5 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	5.768472	2	11.53694	0.0031
Full	25.430946			
Reduced	31.199418			
RSquare (U)	0.1849			
AICc	57.4073			
BIC	62.4755			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	1	1.615179	3.230358
Saturated	3	23.815767	Prob>ChiSq
Fitted	2	25.430946	0.0723

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.79696403	0.4411764	3.26	0.0708
AV[0]	-1.1354868	0.4435489	6.55	0.0105
OP[0]	-0.6845978	0.3919514	3.05	0.0807

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AV	1	1	8.96180596	0.0028
OP	1	1	3.28489346	0.0699

Table 26. Allisions Compared to HFACS Level II Subcategories of Violations (AV) and Organizational Processes (OP)

In testing the third accident type, groundings versus non-groundings across the HFACS Level II subcategories (see Table 27), the analysis revealed that the whole-model test was not useful ($\chi^2_{(14)}$ = 20.35, p = 0.12). The lack-of-fit test revealed that there was no evidence of a lack of fit (p = 0.22). In testing the significance of the predictors in the logistic model, there was a significant difference for the subcategories of Skill-Based Errors (AE1) (p = 0.02), Judgment and Decision-Making Errors (AE2) (p = .01), Inadequate Supervision (SI) (p = 0.04), and Organizational Climate (OC) (p = 0.03).

Nominal Logistic Fit for Grounding
Converged in Gradient, 6 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	10.178865	14	20.35773	0.1193
Full	21.020553			
Reduced	31.199418			
RSquare (U)	0.3263			
AICc	87.0411			
BIC	100.109			
Observations (or Sum Wgts)	48			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	31	18.247964	36.49593
Saturated	45	2.772589	Prob>ChiSq
Fitted	14	21.020553	0.2284

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.3688765	1.3399707	0.08	0.7831
AE1[0]	-1.4078183	0.6207057	5.14	0.0233
AE2[0]	2.22695826	0.8667876	6.60	0.0102
AE3[0]	0.31347605	0.9028895	0.12	0.7284
AV[0]	-0.3499217	0.5613497	0.39	0.5330
PE[0]	0.95495571	0.7859725	1.48	0.2244
PP[0]	0.90429066	0.746379	1.47	0.2257
PC[0]	-0.5906559	0.5101707	1.34	0.2470
SI[0]	-1.1816498	0.5811902	4.13	0.0420
SF[0]	0.17614551	1.0205325	0.03	0.8630
SP[0]	0.03427429	0.5575382	0.00	0.9510
SV[0]	1.17087539	0.9227449	1.61	0.2045
OR[0]	-0.3938131	0.5804898	0.46	0.4975
OC[0]	1.51576926	0.7049128	4.62	0.0315
OP[0]	0.35789385	0.570086	0.39	0.5301

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AE1	1	1	6.83847131	0.0089
AE2	1	1	10.1641933	0.0014
AE3	1	1	0.11824187	0.7309
AV	1	1	0.39781839	0.5282
PE	1	1	1.61992938	0.2031
PP	1	1	1.8719999	0.1712
PC	1	1	1.42373382	0.2328
SI	1	1	5.27642924	0.0216
SF	1	1	0.02959068	0.8634
SP	1	1	0.00377556	0.9510
SV	1	1	1.77395103	0.1829
OR	1	1	0.48216037	0.4874
OC	1	1	5.80036051	0.0160
OP	1	1	0.40729617	0.5233

Table 27. Groundings Compared to All HFACS Level II Subcategories.

The subcategories of Skill-Based Errors (AE1), Judgment and Decision-Making Errors (AE2), Inadequate Supervision (SI), and Organizational Climate (OC) were singled out as the only independent variables and regression analysis was conducted again (see Table 28). The whole-model test revealed that there was evidence to suggest the model was useful ($\chi^2_{(4)}$ = 14.8, p = 0.01). The lack-of-fit test revealed that the model was adequate (p = 0.78). In testing the significance of each subcategory as a predictor in the parameter estimates, the subcategories of Inadequate Supervision (AE2), Inadequate Supervision (SI), and Organizational Climate (OC) were determined to be statistically significant. The subcategory of Skill-Based Errors (AE1) was not considered statistically significant (p = 0.06).

Nominal Logistic Fit for Grounding
Converged in Gradient, 5 iterations
Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	7.408836	4	14.81767	0.0051
Full	23.790582			
Reduced	31.199418			
RSquare (U)	0.2375			
AICc	59.0097			
BIC	66.9372			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	9	2.797562	5.595124
Saturated	13	20.993020	Prob>ChiSq
Fitted	4	23.790582	0.7797

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.56299031	0.504128	1.25	0.2641
AE1[0]	-0.87684	0.4705306	3.47	0.0624
AE2[0]	1.17507748	0.5204014	5.10	0.0239
SI[0]	-1.1368496	0.4925016	5.33	0.0210
OC[0]	1.17851434	0.5939392	3.94	0.0472

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AE1	1	1	4.09535754	0.0430
AE2	1	1	6.59060541	0.0103
SI	1	1	7.26053974	0.0070
OC	1	1	4.58864738	0.0322

Table 28. Groundings Compared to HFACS Level II Subcategories of Skill-Based Errors (AE1), Judgment and Decision-Making Errors (AE2), Inadequate Supervision (SI), and Organizational Climate (OC).

E. MODIFIED HAZARD ANALYSIS

Based on the factors that were identified significant in the logistic regression analysis, a modified hazard analysis process was conducted. This presented prevalent patterns. The modified hazard analysis entailed defining which factors were assessed as significant or prevalent, describing what the effect was for each latent condition, and examining common patterns between them based on evaluating the narratives of the accident reports. Each individual pattern was identified and assessed in terms of risk and composition of severity and probability. Potential intervention strategies currently used in the U.S. Navy were provided as possible mitigations. Table 29 presents the modified hazard analysis.

ACCIDENT TYPE	Pattern Number	ACTIVE FAILURES	LATENT CONDITIONS	EFFECT	Risk Assessment Value	CORRECTIVE MEASURES (MITIGATION)	Risk Assessment Value
Collisions (n = 11)	Collisions Pattern #1	Unsafe Acts	Supervision (S) (n = 12)	Improperly trained and/or inexperienced crew and master leads to lack of procedure compliance	Occasional/Catastrophic (4)	Provide training (BRM/CRM/Shiphandling) as applicable	Remote/Catastrophic (8)
Allisions (n = 15)	Allisions Pattern #1	Used known or unknown faulty equipment	Resource/Acquisition Management (OR) (n = 6)	Ergonomics issues increased workload of bridge team and limited their abilities	Probable/Catastrophic (2)	Develop measures to identify fixed physical hazards	Occasional/Catastrophic (4)
		Failed to assess risk prior to port transit		Risk assessment with control measures not in place	Probable/Catastrophic (2)	Conduct risk assessment and provide risk training	Occasional/Catastrophic (4)
				Process inadequate to identify unknown and known issues	Occasional/Catastrophic (4)	Set in place or evaluate current preventative measures (preventative maintenance plans)	Remote/Catastrophic (8)
	Allisions Pattern #2	Poor communications among bridge team and pilot inhibited proper ship movement	Inadequate Supervision (SI) (n = 7)	Identify ways to lessen the effects of one individual's error on the entire bridge team	Occasional/Catastrophic (4)	Provide training (BRM/CRM/Shiphandling) as applicable	Remote/Catastrophic (8)
		Violation (AV) No supervision resulted in procedures not followed (n = 2)		Bridge team supervision inadequate	Occasional/Catastrophic (4)	Provide training (BRM/CRM/Shiphandling) as applicable	Remote/Catastrophic (8)
				Supervision and communications not effective or not present	Occasional/Catastrophic (4)	Improve communications process and assign qualified supervisor	Remote/Catastrophic (8)
		Supervision failed to assess risk and resulted in unsafe situation		Lack of proactive communications or lack of supervisory response on bridge	Probable/Catastrophic (2)	Improve communications process and assign qualified supervisor	Occasional/Catastrophic (4)
	Allisions Pattern #3	Unsafe actions by operators	Organizational Process (OP) (n = 10)	Lack of procedures and guidance caused unsafe situations for operators	Occasional/Catastrophic (4)	Reevaluate current procedures and checklists. Provide procedural training	Remote/Catastrophic (8)
				Management should have established control measures for identified risk	Occasional/Catastrophic (4)	Conduct risk assessment and provide risk training	Remote/Catastrophic (8)
Groundings (n = 14)	Groundings Pattern #1	Skill-Based Errors (AE1) (n = 9)	Inadequate Supervision (SI) (n = 3)	Procedures/checklists unavailable or not followed correctly	Probable/Catastrophic (2)	Reevaluate current procedures and checklists. Provide procedural training	Occasional/Catastrophic (4)
			Organizational Climate (OC) (n = 4)	Insufficient auditing of checks and/or testing procedures	Probable/Catastrophic (2)	Put in place proper checks or testing procedures for organization	Occasional/Catastrophic (4)
	Groundings Pattern #2	Judgment and Decision-Making Errors (AE2) (n = 14)	Organizational Processes (OP) (n = 13)	Organization did not provide proper risk assessment procedures and control measures	Occasional/Catastrophic (4)	Conduct risk assessment and provide risk training	Remote/Catastrophic (8)
			Inadequate Supervision (SI) (n = 3)	Proper supervision not in place to prevent wrong action	Occasional/Catastrophic (4)	Conduct risk assessment and provide risk training	Remote/Catastrophic (8)
			Organizational Climate (OC) (n = 4)	Insufficient internal auditing of procedures, management, guidance, instructions, and general crew knowledge	Probable/Catastrophic (2)	Reevaluate current procedures and checklists. Provide procedural training	Occasional/Catastrophic (4)

Table 29. Modified Hazard Analysis of Identified Patterns

The following is a characterization of the patterns identified for each type of accident. From the data analysis, it was determined that the HFACS Level I category of Supervision was statistically significant when vessels collided. Further logistic regression using the Supervision category verified this result. Additional regression analysis of the HFACS Level II subcategories of Supervision proved unstable, likely due to the extensive number of variables. The data analysis did reveal one general pattern within the collision accidents. A more in-depth examination of the collision accident reports revealed latent conditions regarding supervisory factors appeared to be followed by unsafe acts. The main effect from this latent condition was that the potential for improperly trained or inexperienced crews tends to result in a lack of procedural compliance. This was assessed as a risk category of four, where the likelihood was considered an occasional occurrence and the severity was catastrophic. Further analysis with a larger sample will likely stabilize the dataset.

Allisions had more significant factors, which provided a sharper representation of identifiable patterns. Three patterns were outlined in the modified hazard analysis. The first pattern observed involved the latent condition of Resource Management, which represented 6 of the 15 accidents. The pattern identified that two unsafe acts could result if the defensive barriers failed. The first unsafe act involved ships' crews utilizing known or unknown faulty equipment. The effect of the latent condition was determined to be human factors issues onboard the ship, creating an excessive workload for bridge teams. This was assessed as a risk category of two, where the likelihood was probable and the severity was catastrophic. The second unsafe act within this pattern was the failure to assess risk while maneuvering in or out of a port. Two effects considered were that risk control measures were not in place and that there was an inadequate process to identify known and unknown issues. The lack of control measures was assessed as probable for the likelihood of occurrence and the severity as catastrophic. The other latent condition was that inadequate processes were in place to identify known or unknown issues. This was assessed as an occasional occurrence with catastrophic consequences.

The second pattern identified in allisions involved Inadequate Supervisory conditions. This represented seven of the allision accidents. The active failures that may result from this latent condition were poor bridge team communications, inhibiting ship movement; lack of supervision, resulting in procedures not followed (Violations); and supervision failing to assess risk, resulting in unsafe situations. The effects from this latent condition ultimately involved the lack of Bridge Resource Management, along with poor supervision and communication. Three of the active failures were assessed as occasional occurrences with catastrophic consequences. The fourth active failure was also assessed as catastrophic, but probable, in occurrence.

The third allision pattern identified was the embedded latent condition of Organizational Processes, which represented 10 of the allision accidents. The effects of this latent condition were the lack of organizational procedures for the operator and no established organizational risk control measures in place. The two effects of the latent condition could result in an unsafe condition and were both assessed as an occasional occurrence, with the severity considered as catastrophic.

There were two patterns identified from grounding accident reports. From the logistic regression analysis conducted, two HFACS Level I categories were significant. Both Skill-Based Errors and Judgment and Decision-Making Errors had latent conditions prevalent in follow-on accident report reviews. Skill-Based Errors represented 9 of the accidents and Judgment and Decision-Making Errors represented all 14 groundings. For Skill-Based Errors, two HFACS Level II subcategories of Inadequate Supervision and Organizational Climate appeared to be latent conditions. Data analysis indicated Inadequate Supervision was significant, but the HFACS Level I category of Organization was near borderline significant ($p = 0.07$). The effects from both unsafe acts, with regards to the respective failed latent conditions, were the lack of procedural compliance;

insufficient testing procedures; improper organizational risk assessment; improper supervision; and improper internal auditing of management, procedures, and crew knowledge. Overall, both patterns were assessed as catastrophic severities. Improper organizational risk assessment and improper supervision were considered occasionally occurring. The remaining three latent conditions were considered probable to occur in time.

The patterns indicated in the modified hazard analysis were all initially assessed for risk, with regards to severity and probability. Mitigating actions and interventions currently used by the Navy were recommended for all patterns with associated latent conditions. When corrective measures were applied, likelihood values were decreased by one level, thereby reducing the overall Risk Assessment Value. Prior to corrective measures, all patterns were assessed as "High" in the Mishap Risk Category. When some latent conditions were removed through mitigation, the risk lessened to "Serious."

F. SUMMARY

For this analysis, there were 48 commercial vessel accident cases appropriated from three countries' accident databases. Cases were coded by two independent raters using the HFACS taxonomy, which was divided into two levels for ease of analysis. After discussion and reconciliation, a consensus was reached and the subsequent agreements were used for data analysis. Cohen's Kappa was used to measure inter-rater reliability for the HFACS Levels between the two raters before and after reconciliation. Logistic regression was used for analysis to determine if patterns were prevalent and if the patterns could be correlated to specific accident types. Statistically significant factors were then placed in a matrix. This aided in the identification of prevalent patterns and was used to construct a modified hazard analysis. The modified hazard analysis refined specific HFACS patterns associated with the three accident types. The patterns found in each accident type were outlined according to the connected latent conditions, active failures, and overall effects. A risk assessment value was specified, along with mitigating actions to reduce overall risk.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSIONS AND RECOMMENDATIONS

A. DISCUSSION

It is recognized that with the reduction of manpower on U.S. Navy ships, there will be a dependency on automated technology to close the gap. In order to gain insight into the challenges this may pose, an analysis was done on civilian merchant ship accidents involving ships utilizing minimal crew. Databases from three countries were consulted and analyzed. Those cases were then reviewed and coded using HFACS, an adaptation of Reason's SCM of human error.

The codes assigned by the raters were verified through a reliability analysis and it was determined that there was adequate agreement to proceed. The initial analysis involved table-displayed data, which subsequently led to a series of logistical analyses. The first round of analysis identified prevalent HFACS categories indicative of the three types of accidents (collisions, allisions, and groundings). This was followed up by subsequent analysis, which focused on specific HFACS Level I categories found to be statistically significant. Further logistic analysis was conducted using known significant categories to identify significant HFACS Level II subcategories. The second round was a full logistic regression analysis conducted on all HFACS Level II subcategories. This attempt did hold some potential, but it was determined that the small sample size and large number of predictors made the analysis too unstable to pursue in this research.

B. RESEARCH QUESTIONS

This research set out to determine if patterns could be identified from the causal factors described in domestic and international commercial maritime accident reports. The research utilized the HFACS taxonomy to systematically investigate these reports in order to answer the following questions:

1. Research Question #1

The use of HFACS provided a taxonomy to organize latent conditions and active failures identified by *post hoc* data analysis. This enabled frequency counts to be developed from the raters' results. The frequency counts enabled the ability to discern generalized patterns for each rater. The results of each rater were then compared and used to identify specific latent conditions or active failures as problematic, by associating the frequently used HFACS codes. The use of chi square analysis for independence revealed that the accident types are not equally likely to be classified as specific HFACS categories. This indicated that the frequencies of the HFACS codes used by the raters were essential to identifying the high-use categories associated with specific accident types.

The process of applying the HFACS taxonomy was particularly useful to discern patterns and prevalent error types. Prevalent latent conditions were recognized due to the frequency of use when the raters coded the accident findings. This essentially set the stage for a more profound analysis of the data using logistic regression techniques. For this research, the data was structured to find further patterns concerning accident types. The HFACS taxonomy, however, facilitates the examination of other types of analysis, such as discerning if patterns are prevalent among systems or environmental factors.

2. Research Question #2

There were a total of six patterns found among the three different types of accidents. Collisions were the only accident type that did not have any significant HFACS Level II categories. There was one HFACS Level I code, *Supervision*, which was considered statistically significant, and its presence was enough to derive a pattern. The accident type, allisions, had three prevalent patterns identified based on the HFACS Level II subcategories. The three patterns found in allisions were identified by the latent conditions within the HFACS Level I categories of *Supervision* and *Organization*. Groundings also had latent conditions and active failures identified within the HFACS Level I categories of *Organization*, *Supervision*, and *Acts*. The similarity between the three accident types was that the latent condition of *Supervision* was significant throughout all accident patterns. Groundings were the only accident type identified to have active failures as the prevalent human error type identified by the raters from the accident reports. This is significantly different from collisions and allisions, where latent conditions were the only factors identified.

The difference found in the accident patterns from collisions, allisions, and groundings were considered subtle, but noteworthy. This research revealed that it may be probable that active failures may be inherent in ship-grounding accidents, and that latent conditions are more likely to be present in collisions and allisions. Further analysis with a larger sample size of accident reports is needed to confirm any validity of this notion. From the data analysis conducted in this research, however, it is reasonable to consider that differences in the patterns can be identified among the three different accident and human error types. This is vital for the human factors studies of maritime accident reports, in order to develop intervention strategies and prevent future maritime accidents from occurring.

3. Research Question #3

The six patterns identified from the three accident types each had a chain of events that included latent conditions. When particular HFACS categories were identified as significant, a review of the accident reports with significant findings was conducted. This was necessary in order to identify how the causal factors aligned in the chain of events leading to the accident. Once this was established, a modified hazard analysis table was developed to better comprehend the patterns and associated hazards. Risk was then associated with the hazards, along with preventative measures.

The strategy of developing a modified hazard analysis table is particularly useful in *post hoc* accident report analyses. This process can be used for developing intervention strategies of latent conditions, thereby reducing the amount of failed defensive barriers and, subsequently, the amount of accidents.

Further analysis with a larger data set would likely reveal more significant patterns, where hazards could be identified and relative risks could be associated. Whether for civilian maritime use or for the Navy fleet, this method is relatively simple to employ and should be considered as a low-cost method of developing intervention strategies.

4. Research Question #4

During the coding process, the raters were instructed to apply one HFACS code for every finding. If they could not fit an appropriate code with a particular finding into the HFACS taxonomy, they were to annotate that finding as *Not Applicable*. The results of the HFACS coding process revealed that out of 518 findings, Rater 1 coded findings from the accident reports as *Not Applicable* 68 times (13%) and Rater 2 88 times (17%). In general, the HFACS taxonomy supported the classification of 450 out of the 518 findings for Rater 1 (87%) and 430 out of the 518 findings for Rater 2 (83%). The reliability of HFACS was further examined by the inter-rater reliability between the two raters. The agreement between the two raters was determined to be good.

Two reasons were cited by the raters regarding the inability to code this many factors. The first reason was that some findings were mechanically related in nature and the raters could not fit the HFACS codes to properly relate to these types of findings. The second reason was that some findings were very specific to the maritime industry and a "best-fit" answer could not be found within the HFACS codes. Instead of attempting to indirectly fit a code to a finding, the most appropriate option was to leave the finding as unclassified. Prior to the raters' conducting their independent coding process, the HFACS taxonomy was reviewed by the raters for applicability. After discussion, consensus was reached and, subsequently, three codes had the words "aircraft/vehicle" replaced with the word "vessel." Six codes were completely disregarded since it was intended solely for the aviation community (e.g., inadequate anti-G straining maneuver). This did not have an effect on the analysis since the HFACS taxonomy was divided into two levels, primarily to alleviate any difficulties with the analysis process. If the HFACS taxonomy were further divided into three levels, there would have been codes intentionally left out of the analysis, but this would likely have had a minimal effect on the results. Since the HFACS taxonomy was initially developed for the aviation community for accident report analysis, it is recommended that an HFACS taxonomy be developed that is tailored specifically to the maritime community. This was evident due to the amount of findings left unclassified by the raters.

C. CONCLUSIONS

The notion of conducting research on maritime accident reports using the HFACS taxonomy was important, since very little human factors research has been applied in a *post hoc* data analysis. The dataset obtained from the three different countries provided enough information to reveal six significant patterns in the three accident types. Further analysis should utilize a larger dataset to stabilize the HFACS Level II data, in order to fit a logistic regression model and reveal more patterns.

Previous research by Lacy (1998), conducted at NPS, has shown that the reliability of HFACS is good and can be applied to the maritime accident reports in a human factors analysis. In order to improve the reliability of HFACS, the taxonomy needs to be relevant to the maritime community. Maritime accident reports are typically written in a generalized manner and not specific to the HFACS taxonomy. This can create issues for the raters when attempting to fit one HFACS code for each accident finding. The governing bodies for transportation from the three countries should adjust their accident reports to consider HFACS analysis, or other forms of human factors analysis, to better determine what latent conditions exists and how to create intervention strategies.

Finally, the Navy should consider using HFACS for additional human factors research on the various classes of ships including the new minimally manned ships entering the Fleet today (e.g., LCS). The OMP project, which ended recently, should have human factors analysis conducted from the last 10 years to determine if any prevalent patterns developed in the U.S. Navy's surface mishaps or near mishaps. The suggested human factors research should not be limited to collisions, allisions, and groundings, but instead should be broadened to include categories where latent conditions may develop into active failures (e.g., personnel injuries). By using the data analysis conducted in this research and developing a modified hazard analysis, the Navy may be able to develop effective intervention strategies, apply it to the Fleet, and reduce the exorbitant costs of mishaps.

APPENDIX. COLLISIONS COMPARED TO ALL HFACS LEVEL II SUBCATEGORIES

Nominal Logistic Fit for Collision
Converged in Gradient, 21 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	28.974587	14	57.94917	<.0001
Full	1.50151e-6			
Reduced	28.974588			
RSquare (U)	1.0000			
AICc	45			
BIC	58.068			
Observations (or Sum Wgts)		48		

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	31	1.50151e-6	0.000003
Saturated	45	0	Prob>ChiSq
Fitted	14	1.50151e-6	1.0000

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>	ChiSq
Intercept	Unstable	649.315579	35141.112	0.00	0.9853
AE1[0]	Unstable	48.2517752	4506.4534	0.00	0.9915
AE2[0]	Unstable	-379.79949	20559.238	0.00	0.9853
AE3[0]	Unstable	-317.71124	20563.077	0.00	0.9877
AV[0]	Unstable	316.295598	16813.297	0.00	0.9850
PE[0]	Unstable	32.619978	6317.7215	0.00	0.9959
PP[0]		-0.1870719	2251.7678	0.00	0.9999
PC[0]	Unstable	158.449846	8836.7307	0.00	0.9857
SI[0]	Unstable	111.082369	6561.9721	0.00	0.9865
SF[0]	Unstable	-30.216709	12406.041	0.00	0.9981
SP[0]	Unstable	31.4100373	2413.6581	0.00	0.9896
SV[0]	Unstable	-237.63858	13807.937	0.00	0.9863
OR[0]	Unstable	110.562144	6852.2344	0.00	0.9871
OC[0]	Unstable	-205.66662	11098.17	0.00	0.9852
OP[0]	Unstable	270.488685	15350.37	0.00	0.9859

For log odds of 0/1

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
AE1	1	1	8.31893961	0.0039
AE2	1	1	30.9235747	<.0001
AE3	1	1	172.634427	<.0001
AV	1	1	31.49627	<.0001
PE	1	1	2.92198e-7	0.9996
PP	1	1	6.78696e-9	0.9999
PC	1	1	17.6689514	<.0001
SI	1	1	16.2787787	<.0001
SF	1	1	1.93294e-6	0.9989
SP	1	1	6.49071682	0.0108
SV	1	1	21.6657626	<.0001
OR	1	1	15.5381958	<.0001
OC	1	1	22.3302316	<.0001
OP	1	1	43.0426371	<.0001

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Airbus. (2005, Dec.). AIRBUS [PowerPoint slides]. Retrieved from http://www.airbus.com/fileadmin/media_gallery/files/safety_library_items/AirbusSafetyLib_FLT_OPS-FOBN-SEQ01.pdf

Alaska Oil Spill Commission. (1990). Spill: The wreck of the Exxon Valdez. Anchorage: Alaska Oil Spill Commission.

Assistant Secretary of the Navy, Financial Management and Comptroller. (2011). Highlights of the department of the Navy FY 2012 budget. Office of Budget: Department of the Navy.

Atlantic Fleet Training and Testing. (2010). Environmental impact statement/overseas environmental impact statement. Retrieved from http://www.aftteis.com

Brantley, M., McFadden, W., & Davis, M. (2002). Expanding the trade space: an analysis of requirements tradeoffs affecting system design (Tutorial). West Point: U.S. Military Academy. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA488461

British Broadcasting Corporation [BBC]. (2012, January 16). Costa Concordia cruise ship captain 'went off course.' Retrieved from BBC: http://www.bbc.co.uk/news/world-europe-16576979

Bowman, W., Crawford, A., & Mehay, S. (2008, September 30). An assessment of the effectiveness of computer-based training for newly commissioned surface warfare division officers (Technical Report). Monterey: Naval Postgraduate School.

Calhoun, S. (2006). Human factors in ship design: Preventing and reducing shipboard operator fatigue. University of Michigan.

Carter, A. (2012, August 1). Prepared testimony. House Armed Services Committee. Washington, D.C.

Cavas, C. P. (2012, January 26). U.S. Navy sidesteps most Pentagon cuts. Retrieved from www.defensenews.com:article:20120126:DEFREG02:301260013:U-S-Navy-Sidesteps-Most-Pentagon-Cuts

Celik, M., & Cebi, S. (2008, September 11). Analytical HFACS for investigating human errors in shipping accidents. *Accident Analysis & Prevention*, 41(2009), 66–75.

Celik, M., & Er, I. D. (2007). Identifying the potential roles of design-based failures on human errors in shipboard operations. *TRANSNAV International Journal on Marine Navigation and Safety of Sea Transportation*, 1(3), 339–343.

Cole, W. (2009, July 7). Hawaii-based ship's grounding detailed. Hawaii. Retrieved from http://the.honoluluadvertiser.com/article/2009/Jul/07/ln/hawaii907070350.html

Commander, Naval Surface Force Public Affairs. (2006, August 11). New LCS training process allows crew members to man ship anytime, anywhere. Retrieved from http://www.navy.mil/search/display.asp?story_id=25034

Commander, Naval Surface Forces. (2010). CNSF mishap analysis. Retrieved from http://www.public.navy.mil/navsafecen/Documents/statistics/ops_research/PDF/10-003.pdf

Curdy, B. (2009, April 20). Cohen's Kappa assessing agreement between two raters (categorical data). Retrieved from http://knol.google.com/k/cohen-s-kappa

Defense Acquisition University [DAU]. (2009, September 15). Human factors engineering. Retrieved from https://acc.dau.mil/CommunityBrowser.aspx?id=314782

Department of Defense. (2005, January 11). Department of Defense human factors analysis and classification system: A mishap investigation and data analysis tool. Washington, D.C.: Department of Defense.

Department of Defense (DoD). (2000, February 10). Department of Defense standard practice for system safety. MIL-STD-882D. Retrieved from https://acc.dau.mil/adl/en-US/255833/file/40632/MIL-STD-882D.pdf

Department of the Navy. (2011, September 1). Department of the Navy implementation and operation of the defense acquisition system and the joint capabilities integration and development system. Secretary of the Navy Instruction [SECNAVINST] 5000.2E.

Dobie, T. G. (2003). *Critical significance of human factors in ship design.* University of New Orleans.

Embry, D. *Understanding human behaviour and error.* Human Reliability Associates. Human Reliability Associates Ltd: Dalton, Lancashire

Galdorisi, G. C., & Truver, S. (2011, July). From minimal to optimal. *Proceedings*, 137(7). Retrieved from http://www.usni.org/magazines/proceedings/2011-07/minimal-optimal

Gates, R. M. (2007). *Zero preventable accidents.* Washington, D.C.: The Secretary of Defense.

Government of Canada. (2012). Transportation Safety Board of Canada. Retrieved from http://www.tsb.gc.ca

Greenert, J. (2011, December). Navy, 2025: Forward warfighters. *Proceedings*, 137(12). Retrieved from http://www.usni.org/magazines/proceedings/2011-12/navy-2025-forward-warfighters

Greenert, J. (2012, March). Statement of Admiral Jonathan Greenert, Chief of Naval Operations before the Congress on FY2013 Department of Navy posture.

Heinrich, H. W. (1941). *Industrial accident prevention: A scientific approach.* New York: McGraw-Hill Book Company, Inc.

Hetherington, C., Flin, R., & Mearns, K. (2006, April 5). Safety in shipping: The human element. *Journal of Safety Research*, 37 (2006), 401–411.

HFACS, Inc. (2010). Resources & Publications. Retrieved from http://hfacs.com/resources/hfacs-publications

Holness, K. S., Shattuck, N. L., Winters, J., Pharmer, J. A., & White, D. (2011, June 6). Human systems integration tradeoff analyses: Lessons learned in support of naval surface acquisitions (Report). Monterey: Naval Postgraduate School.

Joint Services Safety Chiefs [JSSC]. (2012). Policy on the collection and analysis of mishap human factors data. Memorandum of Agreement. Washington, D.C.: Joint Services Safety Chiefs.

Kakesako, G. K. (2011, June 25). Navy ship back at sea after 2009 grounding. Retrieved from http://www.staradvertiser.com/news/201110625_Navy_ship_back_at_sea_after_2009_grounding.html?id=124534899

Kuehmayer, J. R. (2008). AMEM Austrian marine equipment manufacturers*.* Retrieved from http://www.amem.at/pdf/AMEM_MARINE_ACCIDENTS.pdf

Lacy, R. (1998). Human factors analysis of U.S. Navy afloat mishaps (Master's Thesis). Monterey: Naval Postgraduate School.

Lazzaretti, P. C. (2008). HSI in the USN frigate community: Operational readiness and safety as a function of manning levels (Master's Thesis). Monterey: Naval Postgraduate School.

Li, W.-C., & Harris, D. (2006). Pilot error and its relationship with higher organizational levels: HFACS analysis of 523 accidents. *Aviation Space and Environmental Medicine, 77*(10), 1056–1061.

Lin, S.-c. (1998). Physical risk analysis of ship grounding (Master's Thesis). Department of Ocean Engineering, Massachusetts Institute of Technology.

Lippold, K. (2012). U.S.S. Cole. Secretary of the Navy's guest lectures. Monterey: Naval Postgraduate School.

Marine Accident Investigation Branch [MAIB]. (2012). MAIB marine accident investigation branch. Retrieved from http://www.maib.gov.uk

Maurino, D. E., Reason, J., Johnston, N., & Lee, R. B. (1995). *Beyond aviation human factors.* Brookfield, Vermont: Ashgate Publishing Company.

Maybus, R. (2009, July 6). Department of the Navy safety. Washington, D.C.

McPherson, M. (2001). GPS and the joint force commander: Critical asset, critical vulnerability. Newport: Naval War College.

Mulloy, J. (2011, Feb 14). Department of the Navy FY 2012 President's budget [PowerPoint slides]. Office of the Assistant Secretary of the Navy: Financial Management and Comptroller. Retrieved from www.finance.hq.navy.mil/FMB/12pres/DON_PB12_Press_Brief.pdf

Munoz, C. (2012, January 11). Navy to review 313-ship plan, VCNO says. Retrieved from http://www.defense.aol.com

National Resource Council [NRC]. (1994). *Minding the helm.* Washington, D.C.: The National Academies Press.

National Transportation Safety Board [NTSB]. (2012, March 2). National Transportation Safety Board. Retrieved from http://www.ntsb.gov

Naval History & Heritage Command. (2011, June 10). U.S. Navy active ship force levels. Retrieved from ://www.history.navy.mil/branches/org9-4.htm

Naval Safety Center. (2012). Mishap summary. Retrieved from http://www.public.navy.mil/navsafecen/Pages/execsummary/daily_stats_archive.aspx

Naval Surface Forces Public Affairs. (2011, May 20). Afloat Training Group Pacific. Retrieved from file://localhost/www.public.navy.mil/http/::www.public.navy.mil:surfor:atgpac:Pages:1000thGraduateCompletesSWOIntroductionCourse.aspx

NAVSOURCE.org. (2012, March 2). NAVSOURCE archives. Retrieved from http://www.navsource.org/archives/10/16/161194.htm

Oceania Regional Response Team. (2009). Annual report 2009. Retrieved from http://www.oceaniarrt.org/external/content/document/3911/1477659/1/ORRT%20Agenda%20Hawaii%20July%2010%20to%2012%20final.pdf

O'Connor, P. (2008, June). HFACS with an additional layer of granularity: Validity and utility in accident analysis. *Aviation, Space, and Environmental Medicine*, 79 (6), 599–606.

O'Rouke, R. (2011, November). Now hear this – the Navy in a time of less money. *Proceedings*, 137 (11). Retrieved from http://www.usni.org/magazines/proceedings/2011-11/now-hear-navy-time-less-money

Rasmussen, J. (1986). *Information processing and human-machine interaction : an approach to cognitive engineering*. North-Holland: Elsevier Science.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-13 (3), 257–266.

Rasmussen, J. (2003). The role of error in organizing behaviour. *Quality Safe Health Care*, 12 (5), 377–385. doi:10.1136/qhc.12.5.383

Reason, J. (1990). *Human error* [Kindle version]. Cambridge University Press. Retrieved from http://www.amazon.com/Human-Error-ebook/dp/B000SEJ89W

Reason, J. (1997). M*anaging the risks of organizational accidents.* Brookfield, Vermont: Ashgate.

Rothblum, D. M. (2002). Human error and marine safety (Technical Report). U.S. Coast Guard Research & Development Center, Groton, Conneticutt.

Rothblum, D. M., Wheal, D., Withington, S., Shappell, S. A., Wiegmann, D. A., & Boehm, W. (2002, April). Human factors in incident investigation and analysis. Second International Workshop on Human Factors in Offshore Operations (HFW2002), Houston.

Salmon, P., Regan, M., & Johnston, I. (2005). Human error and road transport. Monash University: Accident Research Centre.

Schonberg, J. (2012, January). Maximizing minimum manning. *Proceedings,* 138 (1). Retrieved from http://www.usni.org/magazines/proceedings/2012-01/maximizing-minimum-manning

Shappell, S. A., & Wiegmann, D. A. (2000). The human factors analysis and classification system-HFACS. Office of Aviation Medicine. Springfield: U.S. Department of Transportation Federal Aviation Administration.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. A. (2007, April). Human error and commercial aviation accidents: An analysis using the human factors analysis and classification system. *Human Factors and Ergonomics Society*, 49 (2) 227–242.

Shovlin, K. (2008, September). SWOS-in-a-box. Retrieved from http://www.veteransmagazine.com/membersarea/MagazineIssues/19thmag-sept08/SWOS-in-a-Box.pdf

Surface Warfare Officer School. (2001). Navigation case studies [PowerPoint slides]. Newport, Rhode Island: Surface Warfare Officer School.

Uebersax, J. S. (1987). Diversity of decision-making models and the measurement of inter-rater agreement. *Psychological Bulletin*, 101 (1) 140–146.

United Kingdom. (1995). Merchant Shipping Act 1995. Retrieved from http://www.legislation.gov.uk

United States Government Accountability Office [GAO]. (2010). Military readiness: Navy needs to reassess its metrics and assumptions for ship crewing. Washington, D.C.: GAO.

Whitney, D. (1990, August 1). NTSB spreads the blame for oil spill. *Anchorage Daily News*.

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon Becker, S. E. (2004). *An introduction to human factors engineering.* Upper Saddle River, New Jersey: Pearson Prentice Hall.

Woods, D., Dekker, S., Cook, R., Johannesen, L., Sarter, N. (2010). *Behind human error* [Kindle edition]. Retrieved from http://www.amazon.com/Behind-Human-Error-David-woods/dp/0754678334

Woods, D. (2006, March 28) Human Error with David Woods. *IND ENG 820 at The Ohio State University Podcast.* Podcast retrieved from http://itunes.apple.com/us/podcast/human-error-david-woods-ph.d./id129951355

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. John K. Schmidt
   Naval Postgraduate School
   Monterey, California

4. Samuel E. Buttrey
   Naval Postgraduate School
   Monterey, California

5. Patricia Hamburger
   Naval Sea Systems Command
   Dahlgren, Virginia